January 2010

Unearthing the Roots of Venice: From Relics to DNA

Andrew James Kazanovicz
Worcester Polytechnic Institute

Benjamin M. Allen
Worcester Polytechnic Institute

Debora Afezolli
Worcester Polytechnic Institute

Jaclyn E. Hepworth
Worcester Polytechnic Institute

UNEARTHING THE ROOTS OF VENICE: FROM RELICS TO DNA

AN INTERACTIVE QUALIFYING PROJECT REPORT
WORCESTER POLYTECHNIC INSTITUTE

Submitted to:

Project Advisor: Fabio Carrera, WPI Professor
Project Co-Advisor: Daniel Gibson, WPI Professor

Submitted by:

Debora Afezolli

Benjamin Allen

Jaclyn Hepworth

Andrew J. Kazanovicz

Date: December 18, 2009

http://ve09orig.wordpress.org
ve09-orig@wpi.edu

ABSTRACT

This project sought to contribute to the comprehension of the origins of Venice and its people through archaeology, ancient documents, and genetic genealogy. ArchEasy, a web-based management system for Venetian archaeologists was evaluated and promoted. A promotional website was established for *uScript*, a web-based application that will create a repository of manuscript transcriptions. One hundred Venetian DNA samples were collected in collaboration with the Genographic Project to contribute to the determination of the origins of the first inhabitants of Venetians.

ACKNOWLEDGEMENTS

We would like to thank the following people and organizations for help with our project:

Dr. Marco Bortoletto and Dr. Alberto Zandinella, for taking the time to show us a Venetian archaeology site and explain the intense processes associated with them.

Dr. Giovanni Caniato, for the wonderful tour of the Venice State Archive and support for *uScript*.

Dr. David Comas, for facilitating our collaboration with Universitat de Pompeu Fabra and the Genographic Project, and for giving us a tour of your beautiful biotechnology lab in Barcelona, Spain.

Alberto Gallo, for organizing our apartment and securing our boat passes.

Christine Friesea, for your six-week crash course in Italian. Grazie mille!

John Brunelli (from the Venice 4.0 team), for your technical assistance in developing the *uScript* website, as well as your extensive music and movie libraries that kept us entertained throughout the seven weeks.

Matteo Secchi, Pierluigi Tamburrini, and Venessia.com, for your warm welcome to join the Funeral for Venice theatrics for the collection of DNA samples.

The Settemari, for giving us the opportunity to collect DNA samples during a regatta.

Caffé Brasilia, for allowing us to collect DNA samples in your shop .

President Dennis Berkey, for taking time out of your busy schedule to visit the Venice Project Center.

Professor Fabio Carrera, for selecting each of us for the Venice Project Center for an unforgettable experience, and for your guidance throughout the project.

Professor Daniel Gibson, for your biology expertise, valuable comments on our report, and for assuming the role of photographer at times. A special thanks to Mrs. Gibson for a delicious dinner.

Thanks to the **residents of Ca' Mori, Ca' Corte, and Arco I** for letting us invade your apartments for internet.

And a *special* thanks to **all of the Genographic participants!**

AUTHORSHIP

Each team member contributed equally towards the completion of this project. Without the full participation of each member and extensive teamwork the following project would not have been successful.

TABLE OF CONTENTS

Abstract ii
Acknowledgements iii
Authorship iv
Table of Contents v
Table of Tables vii
Table of Figures vii
Executive Summary ix
1 Introduction 1
2 Archaeology in Venice 4
2.1 Background 4
2.1.1 Contributions of Archaeology in Venice 5
2.1.2 Current Archaeological Process in Venice 8
2.1.3 State-of-the-Art in Archaeology Management 10
2.2 Methodology and Results 12
2.2.1 Digitized Murano Archaeology Site Documentation 12
2.2.2 Redesigned ArchEasy Application 13
2.2.3 Designed a Promotional Website for ArchEasy 15
2.2.4 Drafted ArchEasy Grant Proposal 18
2.3 Conclusions and Recommendations 19
3 Ancient Documents in Venice 21
3.1 Background 21
3.1.1 Contributions of Ancient Documents in Venice 22
3.1.2 Current Manuscript Consultation Process in Venice 24
3.1.3 State-of-the-Art in Manuscript Consultation 25
3.2 Methodology and Results 28
3.2.1 Redesigned the *uScript* Website 28
3.2.2 Researched Funding for the Development of *uScript* 30
3.3 Conclusions and Recommendations 31
4 Genetic Genealogy in Venice 32
4.1 Background 32
4.1.1 Early History of the Veneto 33
4.1.2 Venetian Origins Theories 34
4.1.3 Contributions of Genetic Genealogy in Venice 39

4.1.4 Using Genetics to Solve Origins Mysteries 46
4.2 Methodology and Results 47
4.2.1 Completed the Venetian Pilot Study 48
4.2.2 Identified Future Testing Locations to Validate Origins Theories 49
4.2.3 Established Future Collaboration with the Genographic Project 50
4.3 Conclusions and Recommendations 50
5 Bibliography 52
Glossary I
Appendix A: ArchEasy Grant Proposal III
Appendix B: Screen Shots of ArchEasy Website XVIII
Appendix C: Worcester Polytechnic Institute IRB Form XXI
Appendix D: DNA Consent Form XXX
Informed Consent Form—English XXX
Informed Consent Form—Italian XXXII
Appendix E: Swabbing Instructions & Kit XXXVII
Appendix F: Project Information Card XXXIX
Appendix G: Ancestry Questionnaire XL
Appendix H: Phoenician Scientific Article XLII
Appendix I: Funerale di Venezia LII
Appendix J: Photos from Testing LII

TABLE OF TABLES

Table 1: Important Features of ArchEasy 14

Table 2: Hierarchy of Users and Descriptions XI

Table 3: Forms in the ArchEasy System XII

Table 4: Sample Ancestry Questionnaire XL

TABLE OF FIGURES

Figure 1: Infared Aerial View of Altino 6

Figure 2: Extrapolation of Ancient Altinum 6

Figure 3: Tresso Canal Arrowhead 7

Figure 4: Boat Artifcat Found Near Lova 8

Figure 5: New ArchEasy Organization Scheme 13

Figure 6: Screenshot of ArchEasy Homepage 16

Figure 7: Screenshot of Archeasy's Drop Down Menu 16

Figure 8: Screenshot of ArchEasy Features Page 17

Figure 9: Study Room at Venice State Archive 24

Figure 10: Collection of Documents at Archive 25

Figure 11: *uScript's* Archive Assistant 26

Figure 12: *uScript's* Transcription Assistant 27

Figure 13: Non-CMS Website 28

Figure 14: Previous CMS-Based Website 29

Figure 15: Newly Redesigned *uScript* Homepage 29

Figure 16: Transcript Assistant Animation 30

Figure 17: The Veneto Region of Italy 33

Figure 18: Attila the Hun's Invasion of Italy 34

Figure 19: Paphlagonian Migration Theory 35

Figure 20: Pylaemenes 36
Figure 21: Modern-Day Lusatia 37
Figure 22: Lusatia Theory 37
Figure 23: Brittany Migration Theory 38
Figure 24: DNA Structure 40
Figure 25: DNA Replication Process 41
Figure 26: Y-DNA Genealogy 42
Figure 27: The Y-Chromosome 42
Figure 28: Y-DNA Haplogroups 44
Figure 29: mtDNA Genealogy 45
Figure 30: mtDNA Structure 45
Figure 31: mtDNA Haplogroups 46
Figure 32: Theory Reference Map 50
Figure 33: Rudimentary Example of Advanced Search Function XIV
Figure 34: Screenshot of Example Search Results XV
Figure 35: Intervento Map Example Screenshot XVI
Figure 36: ArchEasy Purpose XVIII
Figure 37: ArchEasy Vision XIX
Figure 38: ArchEasy Demonstration XX
Figure 39: ArchEasy Grants XX
Figure 40. Swabbing Instructions Included in DNA Sample Kit XXXVII
Figure 41. DNA Sample Collection Kits XXXVIII
Figure 42. Front of Information Card XXXIX
Figure 43. Back of Information Card XXXIX
Figure 44: Press Conference Testing LIII
Figure 45: Funeral for Venice Testing LIV

EXECUTIVE SUMMARY

Venice is a city like no other, unique and universally adored by the millions of tourists that visit each year and by the locals who have inhabited the island for generations. The grandeur of the ancient city can still be seen through the modifications made to suit the modern population. Because of this juxtaposition of new and old, the small lagoon is saturated with fascinating archaeological sites and artifacts. Archaeologists analyze these relics in order to contribute to the knowledge of the origins of the inhabitants of Venice. Additionally, there are documents stored at the Venice State Archive that date back to the beginning of the Venetian Republic. Historians spend a great deal of time examining these documents to learn about the ancient city and the people that inhabited it. And there exists many mysteries concerning the origins of the native Venetians. Three prominent theories exist; the Paphlagonia Theory, Lusatia Theory, and Brittany Theory. The theories claim that modern-day Venetians are descendants of distinct ancient populations who migrated to the Veneto from modern-day northern Turkey, modern-day Poland, Germany, and Czech Republic, and modern-day Brittany, respectively.

This study aimed to contribute to the comprehension of the origins of Venice and its people. It was partitioned and implemented through three objectives. This study:

- Improved the archival and visualization of Venetian archaeology
- Promoted and designed an open-source database to extract information from Venetian manuscripts
- Contributed to the determination of the origins of the first inhabitants of Venice

BACKGROUND

Understanding the past is imperative for societies because it defines cultures, traditions, and customs. The origins of a country form its present cultural identity and help it to avoid past mistakes. The city of Venice contains many remnants from past societies. By analyzing these artifacts and determining their histories, archaeologists are able to trace

the roots of Venetians. Although the origins of these people remain unproven theories, the field of archaeology presents one of the ways in which some truth can be revealed. Artifacts in the lagoon have been found that clearly depict Roman influence in Venice in the first century. Similarly, new technologies have been used to analyze land, revealing literal traces in the ground of previous Roman civilizations. Archaeology has even helped to bring to attention the Greek influence on past Venetian populations, and to the populations that inhabited the land as far back as 500 B.C.E. Despite the importance of archaeology in Venice, the archaeological process remains complicated. The Soprintendenza Archeologica per il Veneto mandates various administrative requirements, which are time consuming and manually completed. Relative dating is one of the main ways in which objects are dated. This process requires archaeologists to cross reference each other's finds. Around the world, various systems and technologies have emerged to help archaeologists deal with the great wealth of information. These systems aid with anything from data collection to data management to data interpretation and modeling. Though efforts have been made by various groups to improve the archaeological process in general, none have been directed towards Venice specifically.

The processes within the Venice State Archive are equally inefficient. Approximately ninety (90) kilometers of ancient documents line the Archive shelves, making the searching of desired documents painstakingly difficult. Additionally, transcriptions are made of the same documents many times over since no program is available to facilitate the sharing of documents amongst historians. A program that would allow for the sharing of transcriptions would prove beneficial to historians and the Archive. There are web-based databases employed at various archives around the world, but very few of them are focused upon making a database of updatable transcriptions. A program called *uScript*, however, has been in development for several years now. *uScript* is a web-based application that aims to eliminate the inefficiencies in the current processes employed by the Venice State Archive. It is designed to harness the work of historians to create a digital archive of manuscript images and transcriptions that are accurate and searchable. The system consists of three main components: the Archive Assistant, Transcription Assistant, and Contribution Accountant. The components work together to allow users the ability to search and transcribe manuscripts of interest. These manuscripts are made readily

available to the general public once deposited into the online repository. The Archive Assistant is responsible for the search of the manuscripts and transcriptions and management of the database. The Transcription Assistant aids the user in transcribing the actual document, word-by-word. The last component, the Contribution Accountant, rates the transcriptions and the contributors based on the accuracy of the transcriptions. A great deal of work has been put into the planning and programming of *uScript*, but there is still work to be done to make it a functioning system.

When Attila the Hun and the Germanic Lombards invaded Italy in the 5th Century A.C.E, many Italians sought refuge in the Venetian lagoons. These people are often credited as the first inhabitants of the Venice, however, little is known about the ancient peoples who populated the Veneto in the early centuries B.C.E. Three prominent theories have emerged over hundreds of year that attempt to explain the origins of modern-day Venetians; the Paphlagonia Theory, Lusatia Theory, and Brittany Theory. Based upon ancient documents and archaeology, the theories claim that Venetian ancestors migrated to the Veneto from modern-day northern Turkey, modern-day Poland, Germany, and Czech Republic, and modern-day Brittany, respectively.

To validate the origins of a population, scientists use a technology called genetic genealogy, which uses genetic testing to trace family lineages as far back as thousands of years. Genetic genealogy examines Y-chromosomal and mitochondrial DNA in males for repetitions and unique changes in DNA nucleotide sequences to determine relationships between modern and ancient populations. From these repetitions, called short terminal repeats, and unique changes, called single nucleotide polymorphisms, scientists can determine the migration pattern of one's ancestors, called haplogroups.

In 2005, National Geographic and IBM launched the Genographic Project, which aims to use genetic genealogy to develop a better understanding of human genetic roots. In 2008, a Worcester Polytechnic Institute project group collaborated with the Genographic Project and received three hundred (300) DNA sample kits to conduct a Venetian pilot study.

METHODOLOGY AND RESULTS

In order to improve the archival and visualization of Venetian archaeology, this study:

- Digitized Murano archaeology site documentation
- Redesigned ArchEasy application
- Designed a promotional website for ArchEasy
- Drafted ArchEasy grant proposal

To accomplish this objective, archaeological data from a Murano archaeological site (intervento) was collected, organized, and made electronically available using a scanner and the Google Documents application. This data was also used to create animations explaining the necessity and features of ArchEasy. A visual site map was then created using Microsoft Powerpoint to display the organization of the system and the management of interventii. A promotional website was further produced using a Content Management System to advertise the system and to explain its intent. Lastly, a grant proposal was drafted to procure funds for the development of ArchEasy, and potential grant foundations were selected based on the qualification requirements and amount of funding offered.

In order to promote and design an open-source database to extract information from Venetian manuscripts this study:

- Redesigned the *uScript* website
- Researched Funding for the Development of *uScript*

Because this program has the potential to change the way manuscripts are viewed and utilized, it was necessary to promote the program and gain as much support for its implementation as possible. A promotional website was created, with several design qualifications in mind. It was desired that the web page have a simple, easy-to-use layout, so as not to burden the user. A Content Management System was chosen, because it is easily modified and updated. Additionally, the content needed to explain *uScript* and its various functionalities in full detail, so that interested individuals could get a feel for the system. Pages that used both text and flash animations were added to display the current and proposed processes at the Venice State Archive, in addition to explaining the various

functions of the components. Sections were added to benefit the developers of the program, such as a Research page that links to all pages used to develop *uScript*. In addition to promoting *uScript*, efforts were focused on finding more foundations that could potentially fund the programming of *uScript*. The developers of the program have researched and implemented methods to incorporate cutting edge techniques to enhance *uScript*. Additional programming is required to complete the *uScript* software. The cost of programming has established a need for funding. A grant proposal was submitted to the National Endowment of the Humanities prior to the start of this project. If this bid for funds is unsuccessful, there are other options. Organizations, such as the Venetian Government and the Gladys Krieble Delmas Foundation, could potentially fund the final programming and realization of the *uScript* program.

In order to determine the origins of the inhabitants of Venice, this study:

- Completed the Venetian Pilot Study
- Identified Future Testing Locations to Validate Origins Theories
- Established Future Collaboration with the Genographic Project

This study collected one hundred (100) DNA samples by targeting Venetian organizations that consisted of many male members. Once a participant was identified, a consent form and ancestry questionnaire from were filled out. Next, a buccal (cheek) swab was administered. Lastly, a unique National Geographic Identification Code was placed on an information card for the participant as well as the DNA sample for Genographic record-keeping. Next, scholarly research was conducted to indentify future testing locations to compare with Venetian DNA, in an effort to validate one of the three origins theories. A custom Google map was created and made available online. A journal article was also written, containing evidence for the three major origins theories. Lastly, this study established a future collaboration with the Genographic Project after traveling to Barcelona, Spain and meeting with Genographic Project researchers.

CONCLUSIONS AND RECOMMENDATIONS

This study assessed the Venetian archaeological process and thus determined the requirements for an effective management system, ArchEasy. The necessary features of ArchEasy were planned and described, and efforts were made to promote its development. Future funding for ArchEasy will functionalize all of its features, and incorporate the autonomous agent approach, a technology that will render the system capable of semi-autonomous data interpretation. Once fully developed, ArchEasy hopes to greatly increase the efficiency of the Venetian archaeological process. With the proper technical guidance as well as adequate funding, ArchEasy will become a widely adopted tool by archaeologists, helping them to manage their data and to shine a light on the history of Venice.

This study successfully promoted and designed an open-source database to extract information from Venetian manuscripts. A website was established for the open-source manuscript transcription software called *uScript.* Additionally, research was conducted to identify several major sources of funding for the future development of *uScript*. In the next few years, it would be ideal for funding to be obtained and the programming completed.

This study also successfully contributed to the determination of the first inhabitants of Venice through the collection of one hundred (100) DNA samples, identification of future testing locations, and establishment of a future collaboration with the Genographic Project. Future comparative DNA analysis will validate one of the Venetian origins theories. Following this determination of origins, the publication of a scholarly research article will provide documentation to establish the identity of the origins of Venetians.

1 INTRODUCTION

Over six billion people inhabit the world today, occupying one hundred and ninety five countries, spanning six continents. We share six thousand nine hundred and thirteen spoken languages, nineteen major religions, and one common ancestor. To gain insight into modern day humanity, we must first examine the roots and traditions that define culture and people. Archaeologists devote their lives to examining ancient architecture and relics in an effort to understand long-ago civilizations. Historians, meanwhile, are constantly exploring ancient writings—manuscripts, books, diaries, and letters—in search of the emergence of cultures and new information. Genealogists carefully piece together family trees to determine heritages spanning thousands of years. New tools in each of these fields have been developed which allow for deeper examination of cultures and populations, such as chronometric dating of archaeological artifacts and DNA amplification via polymerase chain reaction. Important information can also be made readily available via the internet, which can reach the far corners of the world. Online databases specific to archaeology and literature have the capability to organize decades of accumulated data, making it accessible to the public. Nowadays, individuals also have the opportunity to use the science of genetics to determine their origins. Online databases, such as Ancestry.com, combine birth and death records with U.S immigration collections and many other features to allow users to create personalized, updated family trees. Such advances in technology help to reveal the forgotten and lost roots of humanity. Technology aids in the understanding of the history and origins of a population, which ultimately is manifested as a unique cultural identity, to be understood and preserved by future generations.

Venice, Italy is a city like no other, where the rich culture and vibrant traditions of the past fuse together with the hustle and bustle of the modern world. But some of the city's greatest treasures and mysteries remain elusive. Architectural artifacts and relics are regularly found, but infrequently publicized. Modern day Venice is built directly above the ancient city, thus the potential of inadvertent finds is high. All public construction, including road and canal maintenance, requires the presence of an archaeologist in an effort to unearth and document these potentially significant artifacts. Information about the ancient lagoon is locked within manuscripts that are threatened with disintegration.

These manuscripts, which occupy ninety (90) kilometers of shelves in the Venice State Archive, hold important information regarding cultural, social, and political deliberations. However, only scholars and experts have the ability to transcribe this information, decreasing its accessibility. And perhaps the most important mystery of all lies within the Venetian people themselves: their origins. The Venetian migratory enigma has long been at the center of controversy; only in recent years has science been able to truly unearth these mysteries that will ultimately preserve the proud and pronounced history and culture of Venice.

Archaeological relics and ancient manuscripts are at the forefront of Venetian history and culture. Archaeological experts have attempted to create centralized archaeological geo-referencing maps that track the locations of the ancient artifacts found beneath the modern city. Unfortunately, these maps are often old or outdated. The implementation of an easy-to-use, updatable, centralized, web-based application will provide researchers with the most current and accurate data. Similarly, web-based applications are being employed to increase the availability of ancient Venetian manuscripts. A program called *uScript* has been established, which aims to digitize the ancient parchment pages of the Venetian manuscripts, making them widely accessible. The program will reveal valuable information about the ancient city which has been forgotten or lost throughout history. The origins of Venice are perhaps the biggest mystery of all. Although little historical evidence exists regarding the origins of Venice directly, historians have developed several theories, which may explain them. These theories outline the migratory pattern of Venetian ancestors. In an effort to trace and validate the migratory path of humankind, National Geographic and IBM have established the National Geographic Genographic Project. The project uses "cutting-edge genetic and computational technologies to analyze historical patterns in DNA from participants around the world to better understand our human genetic roots[1]." A previous Worcester Polytechnic Institute project group collaborated with the Genographic Project by collecting DNA samples from the Venetian population for a Venetian pilot study, in an effort to solve the Venetian

[1] *The Genographic Project - Human Migration, Population Genetics, Maps, DNA.* 1996-2008. https://genographic.nationalgeographic.com/genographic/index.html (accessed October 10, 2009).

migratory enigma. Although recent advances in technology have made it easier to collect and preserve archaeology, manuscripts, and genetic genealogy, these processes can still be refined and optimized.

Strides have been made in archaeology to make the collection, cataloguing, and organization of recorded data user-friendly and easy. In spite of this, few open-source archaeological databases currently exist. ArchEasy, designed by a previous Worcester Polytechnic Institute project group, represents a strong prototype for a future system. Other technological tools that could be implemented alongside the data base are Geographical Information System (GIS). Similarly, *uScript*, an open-source manuscript database, requires a promotional plan prior to its official launch. A strong support staff and funding are also required to further develop and fine-tune its functions. Much work remains in the genealogy realm, even with the advancement of technology in genetics in recent years. A minimum of one hundred of the three hundred kits that were donated to the previous group must be used to provide a sufficient base for genetic analysis of the Venetian Pilot Study. These local DNA samples will be integrated into the Genographic Project alongside the thousands of other samples collected from all over the world. A future collaboration was established to trace the origins of Venice through the collection and analysis of additional DNA samples in regions where Venetian ancestors are believed to have emigrated.

Refining current methods and processes in archaeology, manuscripts, and genetic genealogy will aid in the preservation of Venetian history and culture. To improve the accessibility of Venetian archaeology, the study aimed to evaluate and promote ArchEasy. A grant shell was drafted in an attempt to procure funding through the National Endowment for Humanities foundation. To refine and promote *uScript*, this study aimed to create a sophisticated promotional web page to ready the program for its official launch. To contribute to the discovery of the origins of Venice, this study aimed to collect over one hundred total DNA sample kits, and lay the foundation for a future collaboration, and to write a scholarly journal article regarding the origins of Venetians based on the prominent preexisting theories.

2 ARCHAEOLOGY IN VENICE

The jewel of the Mediterranean, Venice, has a rich history as a significant political power and a traveler's delight due to its distinctive canals and gorgeous architecture. Despite the city's small size, the field of archaeology in Venice continues to thrive and occupy the attention of city officials. This project aimed to promote and develop a web-management system for archaeological data, to aid in its organization, geo-referencing, and ultimately interpretation. A background on Venetian archaeology is presented in section 2.1, with the contribution of archaeology to Venetian history, the current archaeological process in Venice, and a summary of the state-of-the-art in archaeological management organized as subsections 2.1.1, 2.1.2, and 2.1.3 respectively. The methods used and results obtained can be seen in section 2.2, with the sub objectives organized in subsections ranging from 2.2.1-2.2.4. Lastly, conclusions and recommendations for the future of this project are presented in section 2.3.

2.1 BACKGROUND

The study of archaeology greatly contributes to the understanding of human roots. What the earth has covered up and the seas have swallowed is an essential piece of humanity. People strive to understand their past in order to shape their cultural identities and to avoid repeating mistakes. Venice in particular is a unique city with many archaeological riches, but lacks an organized system of storing and visualizing data. Examples of Venetian archaeology and their contribution to the understanding of Venetian origins are first outlined in section 2.1.1. The current archaeological process in Venice is then described in section 2.1.2, followed by a brief overview of current state-of-the-art in archaeological data management in section 2.1.3.

2.1.1 CONTRIBUTIONS OF ARCHAEOLOGY IN VENICE

Archaeology is the study of the human past by examining material evidence, such as artifacts and data[2]. As described by archaeologist and professor, Martin Carver, "a knowledge, or at least an awareness, of the past is crucial in order to understand modern society through the evaluation of past social interaction[3]. The Veneto region of Italy is a small area saturated with fascinating archaeological sites and artifacts. However, a dilemma is posed by the existence of the modern city of Venice while an ancient world of treasures lies beneath. The origins of Venice's inhabitants remain unclear, but archaeologists have provided important clues to help solve this mystery. A mosaic Roman villa floor was discovered by archaeologist Ernesto Canal in 1991 three meters underwater on the small island in the northeastern part of the lagoon known as San Lorenzo di Ammianai[4]. In addition, over two hundred submerged structures remaining from Roman times were discovered scattered from Lio Piccolo into the swamp of Centrega. These structures are sufficient evidence to support the theory that Romans founded the city of Venice in the first century.

The ancient Roman city of Altinum has recently grabbed the attention of scholars studying the origins of Venice. Its position at the edge of the Venetian lagoon and at the crossroads of Via Annia and Via Claudia Augusta made it a significant commercial center. The Romans actually pioneered the engineering used for the Venetian canals in Altinum, intended to drain the marshes and carry water away from the habitations. After attack from Attila the Hun, the Romans had to flee to the islands of the lagoon and to what is now modern Venice, abandoning Altinum. The marshes took over and the canals filled in, so that

---

[2] White, Nancy. Types of archaeological data. 2004 [cited September 28 2009]. Available from http://www.indiana.edu/~arch/saa/matrix/ia/ia03_mod_05.html (accessed October 3, 2009).

[3] Carver, M. 1996. On archaeological value. *Antiquity*(70): 45 (accessed September 28, 2009).

[4] Pulliero, Augusto. 2002. In laguna nord alla ricerca delle "radici". *Il Gazzetino Di Venezia*, http://www.archeosub.it/articoli/laguna/ammiana.htm (accessed September 18, 2009).

the ancient city literally sank and is now known as Altino[5]. Italian geomorphologists surveyed Altino's fields with infrared cameras during a 2007 drought, and inadvertently discovered remnants of Altinum. Several streets, a basilica, an amphitheatre, and even a canal are clearly visible in their photograph and interpretation, seen in Figure 1 and 2 respectively, providing more evidence for the Roman origins of Venice.

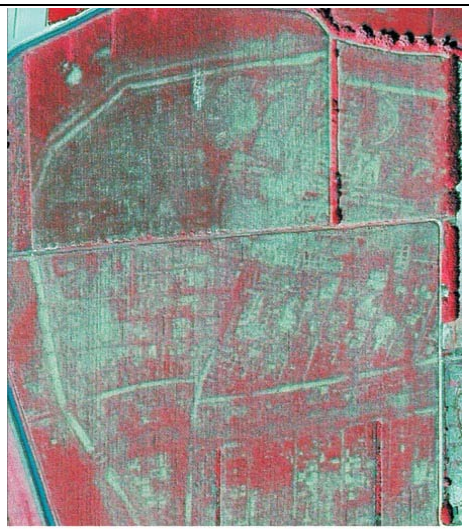


FIGURE 1: INFARED AERIAL VIEW OF ALTINO[6]

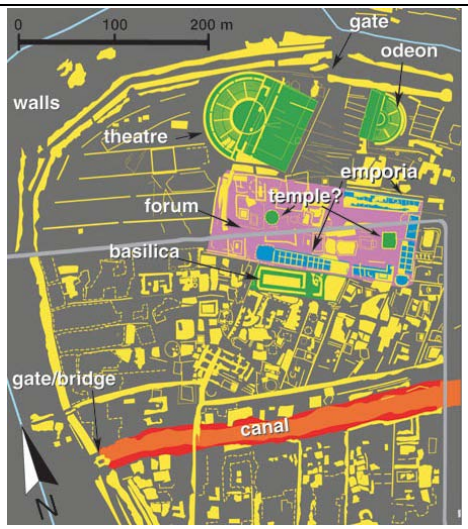


FIGURE 2: EXTRAPOLATION OF ANCIENT ALTINUM[7]

[5] Neal, C. 2009. *Beneath the corn, the husks of ancient venice*. The vigorous north. Vol. 2009, http://vigorousnorth.blogspot.com/2009/08/beneath-corn-husks-of-ancient-venice.html (accessed October 2, 2009).

[6] Image courtesy of http://vigorousnorth.blogspot.com

[7] Image courtesy of http://vigorousnorth.blogspot.com

Modern day archaeologists have made discoveries in the lagoon that also demonstrate a Greek presence, a thousand years before that of the Romans. An ancient Greek pottery head was found in the island of Mazzorbo. Whole pieces of Greek pottery, mostly kylix (wine cups) and skiphoi (cups) dating back to the fifth and fourth centuries B.C., were also found at Altino. However, these were likely supplied by Estruscans of the time period, not the Greeks, and so Venetian roots remain unclear.

FIGURE 3: TRESSO CANAL ARROWHEAD[8]

It is believed that societies existed in Venice even before the Romans or Greeks populated the region[9]. Proof of an environment rich in fresh water and vegetation, as well as certain archaeological discoveries, all lend truth to this theory. However, the artifacts are buried up to four meters deep and often underwater. Discoveries thus occur sporadically and data is not analyzed collectively. An example of such a discovery was an arrowhead discovered in the soil of the Tresso canal, seen in Figure 3, near the island of Lazzaretto, confirming the existence of hunting and fishing thousands of years ago.

A canoe carved from a large oak tree was also found in 1983 at Lova, buried 2.45 meters deep, seen in Figure 4. In 1999, this canoe was dated with Carbon 14 and dendrochronology to the Middle Ages, around 1004 to 1032 A.D[10].

[8] Rossi, Antonio. 1984. La laguna di venezia prima dei romani. Archeologia Viva(March), http://www.archeosub.it/articoli/laguna/lgnprst.htm (accessed September 18, 2009).

[9] Rossi, Antonio. 1984. La laguna di venezia prima dei romani. *Archeologia Viva*(March), http://www.archeosub.it/articoli/laguna/lgnprst.htm (accessed September 18, 2009).

[10] Martinelli, Nicoletta, and Olivia Pignatelli. 1998. Datazione assoluta della piroga di lova (venezia). *Bollettino Del Museo Civico Di Storia Naturale Di Venezia* 49, : 207--212, http://www.archeosub.it/articoli/imbarcazioni/lovamonox.htm (accessed September 18, 2009).

In addition, wood foundations were found underwater near Burano in the northern lagoon that can be traced back to 500 B.C. using Carbon 14 dating. These sparse findings hint at the life that inhabited the lagoon thousands of years ago, and will only clarify the origins of the ancient Venetians once examined collectively.

FIGURE 4: BOAT ARTIFCAT FOUND NEAR LOVA[11]

2.1.2 CURRENT ARCHAEOLOGICAL PROCESS IN VENICE

In the Venetian lagoon, underwater archaeology is as prevalent as land archaeology. Although the discovery and documentation techniques used in underwater archaeology are similar to the ones in land archaeology, special skill sets are required of underwater archaeologists. In Venice, the ability to unearth artifacts goes hand in hand with the ability to maneuver through water. For example, in the Mediterranean, helmeted sponge divers are responsible for having made some of the most important archaeological discoveries[12]. Obtaining data is the first hurdle that must be overcome, followed by the crucial interpretation of this data.

Radioactive carbon dating is a technique that detects the presence of the isotope carbon-14, which can be found in bones, wood, or ash found in archaeological sites[13]. The rate of activity decay of this isotope is known and can be compared to the known value for living organic matter to determine the age of an object. The relative dating principle, also

---

[11] Martinelli, Nicoletta, and Olivia Pignatelli. 1998. Datazione assoluta della piroga di lova (venezia). *Bollettino Del Museo Civico Di Storia Naturale Di Venezia* 49, : 207--212, http://www.archeosub.it/articoli/imbarcazioni/lovamonox.htm (accessed September 18, 2009).

[12] Archaeology. in Encyclopædia Britannica [database online]. 2009b [cited October 6 2009]. Available from http://www.britannica.com/EBchecked/topic/32572/archaeology.

[13] Archaeology. in Encyclopædia Britannica [database online]. 2009b [cited October 6 2009]. Available from http://www.britannica.com/EBchecked/topic/32572/archaeology.

known as cross dating, allows archaeologists to determine the age of undated artifacts by noting which dated objects they are in contact with. Stratigraphy[14], the study of rock successions in relation to time, is used by archaeologists to analyze the layers of rock that have accumulated on sites and to subsequently extrapolate chronology. Because the Veneto region is in a coastal location, this type of cross-referencing is made more difficult by the variation in sea levels. As the availability of records increases, so does the ease with which artifacts can be dated using the aforementioned techniques.

Upon the start of an intervento, or excavation project, Venetian archaeologists must follow the guidelines set forth by the Soprintendenza Archeologica per il Veneto, Nausicaa. An excavation begins by annotating the locality and civil number, the director of the project, the company executing the project, the chief of the building site, and all other information required by various forms. Archaeologists then proceed to photograph the general context of the entrance and panoramic views of the building. Subsequently, the general context of the room or excavation area must be photographed. Lastly, the excavation area must be photographed with metric references, with arrows on a blackboard from the North. All photographs should be taken with a minimum three exposures. The archaeologist has to then create a plan containing the area of the intervento to eventually hand it, and to also agree on the terms of the excavation with the head of the project[15].

The information mentioned above as well as further details are filled out in the Scheda di Intervento Archeologico, or the archaeological intervento form. This encompasses the administrative information that is required from every intervento. The Soprintendenza mandates that a Unita Stratigrafica form (US) be filled out for every layer that is unearthed. These forms are arbitrarily numbered. There is also a similar USM form for objects discovered above ground, although the US forms are more common. Layers are distinguished by their color, components and consistency. The three components of soil are

[14] Stratigraphy. in Encyclopædia Britannica [database online]. 2009d [cited October 6 2009]. Available from http://www.britannica.com/EBchecked/topic/568372/stratigraphy.

[15] Bortoletto, Marco. Manuale di Procedura Archeologica. 2007, pg. 1.

sand, clay, and silt, and the type of material is wholly dependent on the grain size. Archaeologists sieve the material to determine size, and thus can differentiate between layers.

The archaeologists, using a pre-specified numbering system determined by the Soprintendenza Archeologica, then fill out larger index forms by hand. This numbering system is used by the entire country to organize the various archaeological finds, and is the one referenced by the Soprintendenza. Lastly, for every object and layer recorded by the archaeologist, an expert from the Soprintendenza completes an RA form, which also uses the pre-specified number.

Dating in Venice is often accomplished by cross-referencing other objects, and also by placing finds in historical context. If two walls are found to cross each other but do not weave together, then it can be determined that these walls are from different time periods. If these walls are in fact woven, then the construction details must be studied to determine if they are of the same or different epochs. This includes the size of the bricks, the material, and the mortar itself. Similarly, in dating ceramic fragments, the fill is often used to determine the age in which it was produced. It is known that certain fragments were only created during particular time periods, thus narrowing the possibilities of their age. Historical events give power to the relative dating system. The more facts and objects correlate, the easier it becomes to accurately date them. Herein lies the need for a web management system to allow archaeologists to complete the administrative requirements as well as to organize their data.

2.1.3 STATE-OF-THE-ART IN ARCHAEOLOGY MANAGEMENT

When data is made available through one predominant location, it can be collectively analyzed and modeled. Tools such as GIS, or Geographic Information Systems, have become extremely useful to the archaeological process. Archaeologists typically use GIS and graphics for rescue archaeology, large site excavations, regional settlements studies and Cultural Resources Management (C.R.M.)[16]. Different GIS applications are

[16] Djindjian, F. 1998. GIS usage in worldwide archaeology. *Archeologia e Calcolatori* (January): 19--30, http://soi.cnr.it/archcalc/indice/PDF9/09_03_Djindjian.pdf (accessed September 28, 2009).

available for the varying needs of its users, with the main goal of managing various graphic objects. Archaeologists can build a GIS site that allows direct interaction with the excavation data, management of this data, and also provides 3D modeling tools

Various archaeological web management systems are available from different areas of the world. Open Context[17] is one such system that allows its users to input primary field research data from archaeology and related fields. By making this data easily available electronically, it can be found and reused in the future by others without any hassle. Open Context hosts many different projects, with different methodologies, recording systems, and temporal and geographic coverage. It also uses many web-services so that data can be transferred back and forth between Open Context and other applications. Perhaps a very unique aspect of this service is the way in which it facilitates information retrieval by the use of Archaeological Markup Language (ArchaeoML). This system uses an item-based information model, where individual atomic units of observation are related to each other and their descriptions. This allows for deeper searching, with a variety of linked relationships displayed, not only the obvious. ArhcaeoML is one of the technologies that may be adapted to ArchEasy.

The Online Cultural Heritage Research Environment (OCHRE) is another such system, specifically an Internet database system for cultural heritage information started at the University of Chicago by David Schloen. The most unique feature of this system is its ability to integrate data of diverse origins by organizing into a coherent and searchable body of information[18]. The administrator of a particular project can import data tables into the OCHRE system and also export data from the central database at any time. OCHRE uses the XML Query Language to search and manipulate the data stored in the central database, which is preferable when working with hierarchical semi-structured data used in cultural heritage research, such as archaeology. Each XML document in the database corresponds to a real-world entity such as an archaeological site, context, or artifact, etc. These items are

[17] Open context: Community-based data sharing and tagging. in Alexandria Archive Institute; UC Berkeley School of Information [database online]. b [cited September 18, 2009 2009]. Available from http://opencontext.org/.

[18] Schloen, David. OCHRE: Online Cultural Heritage Research Environment. University of Chicago. Accessed December 2, 2009. Available from http://ochre.lib.uchicago.edu/index.htm

grouped into several general categories such as “Locations & Objects”, in order to provide the most efficient storage and retrieval of the information. Data is linked in this way, and can be integrated from a single project to many different ones. The item-based data integration employed by the OCHRE system could be adapted to ArchEasy to allow archaeologists to store and use their data most effectively.

2.2 METHODOLOGY AND RESULTS

Improving the archival and visualization of Venetian archaeology was accomplished using a series of sub-objectives, described below. The first step was to digitize the documents obtained from a Murano intervento in an effort to better understand the administrative process of an archaeological dig, explained in section 2.2.1. The next step, detailed in section 2.2.2, was to utilize this information to redesign the existing ArchEasy system to make it more organized and detailed, so as to be effectively understood by a future programmer. All the information available was then incorporated to create a promotional website for the system, as seen in section 2.2.3. The first draft of a grant proposal was also created to procure funds for the future development of the system, explained in section 2.2.4. These efforts strive to progress ArchEasy to the point where it can aid archaeologists in their search of the history and origins of the Venetian people.

2.2.1 DIGITIZED MURANO ARCHAEOLOGY SITE DOCUMENTATION

The drawings and maps created by archaeologists at the Museo Vetrero Murano intervento were collected from Marco Bortoletto’s and Alberto Zandinella’s archives. Additionally, Dr. Bortoletto’s archaeological procedural binder containing various forms and instructions was also collected.

All of the collected information was scanned using an HP Photosmart C4380 printer and scanner. The documents were scanned and organized into corresponding folders, namely US forms, Schedule forms, Photographs, or Coversheets for the Murano intervento documents, and simply a Procedural binder folder for those sheets. Certain folders also contained subsections. For example, the US forms folder contained completed US forms arranged in subfolders according to their numbering systems. The larger maps and

diagrams were taken to a scanning and photocopying shop, and were scanned onto a flash drive. These maps and diagrams will eventually be converted into GIS to comprise the different layers of the SandTable for future demonstrations. Each stratigraphy (layer of land) and object of interest found within these records will be inserted in the program as a separate layer, to be displayed at the proper time. All of the documents in their appropriate folders were then uploaded into a Google Doc account and subsequently linked on the Origins B2009 project folder, for use by future project groups.

2.2.2 REDESIGNED ARCHEASY APPLICATION

After researching existing archaeological management systems as well as the archaeological process in the city of Venice, various design changes were made to the ArchEasy system to improve its organization, increase its level of detail, and best demonstrate the needs of its users. This was done by creating a mockup of the ArchEasy system website using Microsoft PowerPoint. The site was modeled in design after the Nabonidus Archaeology web management system *http://www.nabonidus.org*. A future programmer can then implement these changes by using the PowerPoint as a guide. The new website theme and organization was modeled after the simple display theme used by Google. Actual intervento data was used to simulate an archaeological map on the system website using the Google Map software. A screenshot of the new organization scheme created by this group can be seen in Figure 5.

FIGURE 5: NEW ARCHEASY ORGANIZATION SCHEME

It was determined that ArchEasy has four major features that were highlighted and expanded upon in the redesigned website, which are explained below in Table 1. Although

these are not exclusive, they were chosen because they give a general idea of its power as an archaeological management system.

Feature	Highlights
Forms	Administrative requirements of the Soprintendenza Archeologica, linked to interventii/sites, automatically saved to database
Personalized Workspace	Log in, create new intervento or continue old intervento, set permissions for all information, hierarchy of users
Database	Generated from forms and uploaded information, searchable on multiple levels
Data Modeling	Interactive archaeological potential map, autonomous agent approach

TABLE 1: IMPORTANT FEATURES OF ARCHEASY

Archaeologists using the system will first be prompted to enter their login information on the program website, after which point they can freely browse the site. They can choose to start a new intervento, or to work with an old intervento. If starting a new intervento, archaeologists will fill out the appropriate forms according to the site and to the type of find. They set permissions for these, from highly restricted to open to the general public. Images, drawings, and GIS files can also be scanned or uploaded into the system. This particular information will eventually be linked to an interactive map that displays the most updated interventi data in a visual and friendly interface.

In future developmental stages, an autonomous agent approach will be employed in which each GIS representation of an object and layer will be capable of reasoning for itself. In this way, objects correlate will with one another without the need of a greater directing entity. This will, in essence, achieve a level of automatic referencing that is unparalleled in other archaeological web management systems. This new technology is not intended to replace the judgment of the experts, for archaeology is an interpretative field that requires

careful consideration. However, it will act as a tool in proposing possible relationships, to then be more carefully examined by archaeologists.

The users of the ArchEasy system can also use the search function to extract the desired information from the database. Searching is made possible by tags and keywords, to be attached to each object or strato (layer) by the archaeologist when filling out the forms. An advanced search option was included in the mockup, in addition to simple word searches.

A hierarchy of users was created in order to create the most effective system of sharing information. This can be seen in the ArchEasy grant proposal in Appendix A. The necessary forms required by the Soprintendenza Archeologica per il Veneto are already displayed online, and have been linked to directly from the PowerPoint demonstration.

2.2.3 DESIGNED A PROMOTIONAL WEBSITE FOR ARCHEASY

ArchEasy has laid the foundations for a future functioning database system, and a promotional website has been created using WordPress to showcase all of the program's capabilities. This can be found at *http://veniceprojectcenter.org/~origins*. WordPress was chosen because it is a Content Management System site, and it is relatively easy to learn and maintain. The site was modeled in content after the *uScript* promotional website. The intent of the website is to explain why ArchEasy is necessary in the city of Venice, what its vision is, and what features have been designed to help it achieve this vision. Additionally, interested archaeologists will eventually be able to request a user account directly from this website. The ArchEasy system demonstration site is accessible from this location, and Professor Fabio Carrera is provided as the contact. Animations, images, and other devices were used to draw attention to important facts. Information was presented in small quantities, to be easily scanned and understood.

A screenshot of the ArchEasy homepage can be seen below in Figure 6. The top menu bar will constantly be visible, and contains the following sections: Home, ArchEasy, Grants, Register, and Contact Us.

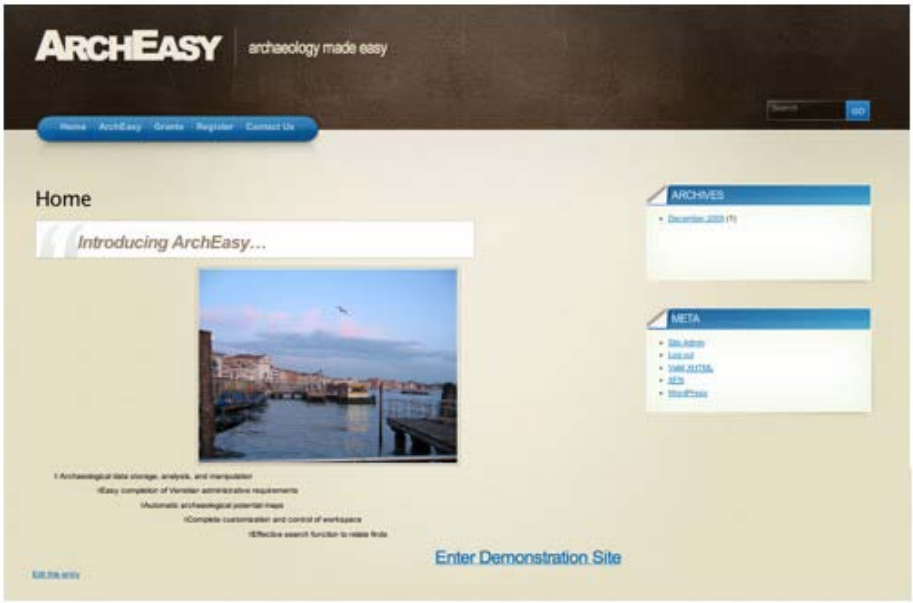


FIGURE 6: SCREENSHOT OF ARCHEASY HOMEPAGE

The Home link will bring the user back to the above page, which will link to the ArchEasy system demonstration website by the "Enter Demonstration Site" link. The ArchEasy menu bar is further split into Features, Purpose, Vision, and Demonstration Site, as can be seen in Figure 7 below.

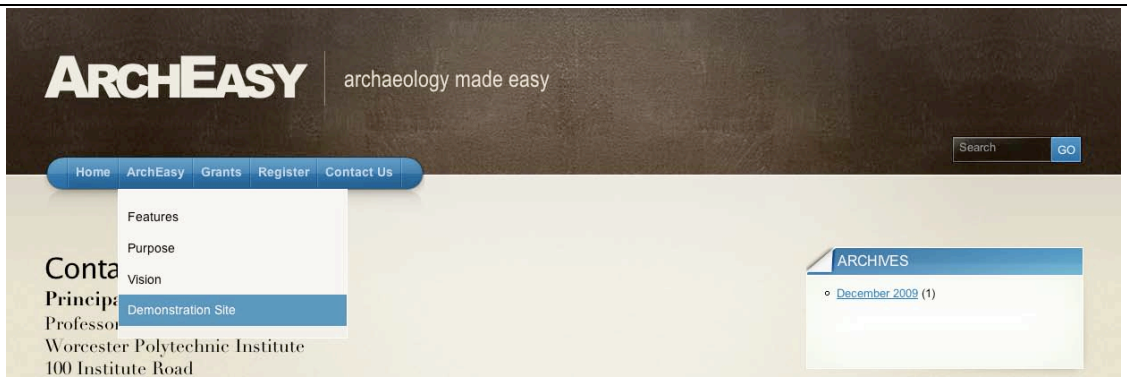


FIGURE 7: SCREENSHOT OF ARCHEASY'S DROP DOWN MENU

The features page presents a concise outline of the four components of the system and can be seen as Figure 8 below.

FIGURE 8: SCREENSHOT OF ARCHEASY FEATURES PAGE

The purpose page presents an explanation of why ArchEasy is necessary in the city of Venice, while the vision page a brief summary of what future functionalities will make ArchEasy appealing to Venetian archaeologists. The demonstration site page provides yet another link to the ArchEasy system demonstration website. The grants page provides a link to the preliminary grant proposal written by this project team for the development of ArchEasy. Screenshot images of each of these pages can be seen in Appendix B. Lastly, the register page will eventually allow interested users to submit their personal information to the ArchEasy administrator in order to obtain an account, or to simply submit their e-mail address to receive news updates. Similarly, the contact page contains the contact information of the principal investigator, Dr. Fabio Carrera, and of the Venice Project Center.

2.2.4 DRAFTED ARCHEASY GRANT PROPOSAL

In order to obtain the necessary funding for the development of ArchEasy into a fully functioning system, a preliminary grant proposal was written, and can be found in Appendix A. This grant proposal contains all of the basic information about the ArchEasy program, as well as more specific information regarding the autonomous agent technology. To write this proposal, research was first conducted on the archaeological process in Venice as well as the field of archaeology in general. This draft of the grant proposal focuses on background information, benefits, and components of ArchEasy. The highlighted program features included the autonomous agent technology, as well as the easily searchable and accessible archival storage of archeological data. This proposal was modeled after the *uScript* National Endowment for the Humanities grant created by Professor Fabio Carrera.

Researching the different grant options was an important step. The possible grant foundations underwent a careful screening process to ensure that they provided what is needed to finish the development and implementation of the program. The criteria that each grant had to meet is listed below:

1. The grant must be applicable to the ArchEasy project and the improvement of digital humanities. This ensures that the final grant(s) that is(are) applied for will actually provide the funding, and will reject the proposal based solely incompatibility.
2. The grant must contain upwards of $15,000. This ensures that the grant money will cover the initial cost of programming the system and the development of autonomous agents technology.
3. The due date must, at the earliest, be February 15th 2010. This will allow the grant to be properly written and critiqued, as well as allow adequate time for revisions.

Various grant resources were employed when researching the grant opportunities. Sources such as previous grants on similar topics, the WPI Research Librarians, and grant sourcing

agents like grants.gov were used. Four grants were ultimately selected, and subsequently reevaluated based on a new set of criteria in order to rank them accordingly. These grants were ranked in the following manner:

- **The best overall start-up grant:** Digital Humanities Start Up Grant
- **The best overall money grant:** Institutes for Advanced Topics in Digital Humanities
- **First Runner Up:** Gladys Krieble Delmas Foundation
- **Second Runner Up:** Genographic Legacy Fund

The Digital Humanities Start Up Grant was the best overall start up option. It offers up to $50,000, and chances of obtaining it are higher because forty (40) start-up grants are distributed. The best grant option for obtaining the most funding was the Institutes for Advanced Topics in Digital Humanities. It offers up to $250,000 for the advancement of programs to preserve digital humanities, what ArchEasy aims to achieve. The first runner up grant is the Gladys Krieble Delmas Foundation, which offers $19,900 for the use of technology to improve current historical information systems. The second runner up grant is Genographic Legacy Fund. In order to obtain it, careful wording of the proposal would be required. However, the amount of funding offered in negotiable, with no set limit or timeframe for project completion.

2.3 CONCLUSIONS AND RECOMMENDATIONS

This team first assessed the Venetian archaeological process by speaking with archaeologists and analyzing intervento data. Next, the features of ArchEasy , a web-management system for Venetian archaeological data, were enhanced. Additionally, a promotional website was created to stress the need for such a system, as well as to highlight its components. Various areas of the previously existing ArchEasy system demonstration website were also updated in an effort to make them clearer to potential users. Lastly, a grant proposal draft was created for the future acquisition of funding. The potential of ArchEasy to greatly increase the efficiency of the Venetian archaeological

process was outlined in the proposal. The autonomous agent approach was also explained, a feature of ArchEasy that would render the system capable of semi-autonomous data interpretation.

As novice computer users, the group has taken the design and plan of ArchEasy as far as possible. Its purpose and vision have been thoroughly explored and explained. What remains is for skilled computer programmers to implement the designs into a fully functioning system. This will not be a simple task, as the components of ArchEasy are complicated and some of the future features, for example the autonomous agent approach require technology that has not been fully developed. Once the system is programmed, the last step will be to present the program to archaeological institutions and companies so that ArchEasy becomes widely adopted. The greater the amount of data stored in its database, the better the quality of data interpretation will be. This is the key to using archaeology to study the past: cross-dating finds with one another. With the proper guidance and effort of experts, the group hopes that ArchEasy will become archaeologist's Swiss army knife, helping them to manage their data and shining a light on the history of Venice.

3 ANCIENT DOCUMENTS IN VENICE

In its heyday, Venice was a formidable power in the economic and political realms. Because of this, a significant number of documents were created, outlining everything from day to day duties of the leaders of the Republic, to letters from foreign diplomats, to massive hand-drawn maps showing the world they knew. All of these documents have been saved and archived in Venice, at the Venice State Archive. This project aims to optimize the consultation of these manuscripts within the Archive, through the use of revolutionary technology and the current work of historians and scholars within Venice. It also aims to educate interested individuals of this cutting-edge technology, to gain support for its implementation. A background on archives and their methods for the consultation of manuscripts is presented in section 3.1. The methods used to promote the *uScript* system, and the final results obtained are shown in section 3.2. The last section, 3.3, outlines the conclusions drawn from the study, and the recommendations for the future of *uScript*.

3.1 BACKGROUND

Ancient writings are one important method to study the history of a people or a place. Venice is fortunate to have a relatively unbroken record of its history in the form of manuscripts. These manuscripts have the potential to aid historians in their research and to aid the city in understanding its origins. Similar to the Venetian effort, many organizations worldwide have been working to utilize the information that can be found in ancient manuscripts. Examples of these efforts are detailed 3.1.1. Unfortunately, these manuscripts are not accessible in a manner that allows for this information to be shared effectively. This is shown in section 3.1.2. There is, however, an alternative to the current methods of studying the manuscripts of ancient Venice that utilizes the forefront of today's technology. This alternative is outlined in section 3.1.3.

3.1.1 CONTRIBUTIONS OF ANCIENT DOCUMENTS IN VENICE

Since manuscripts are extremely valuable to understanding history and culture, historians and scholars work to extract as much information from them as possible, while attempting to preserve them for future reference. The latest method of preserving these documents is by digitizing them and making them available for perusal via computer databases and the Internet. In June 2009, a four (4) day workshop was held at the Radyapustaka Museum in Solo, Indonesia to digitize the museum's entire collection of early documents. Experts from all over the world participated in teams of four (4) to clean, catalogue, and photograph each page. These pages are now available for perusal, via e-library programs[19]. Similarly, Switzerland has instituted a project called "e-codices – Virtual Manuscript Library of Switzerland." The goal of this project was to "provide access to medieval and selected early modern manuscripts held in Switzerland." Countless pages were made available to the public via their website, organized in an easy, searchable manner[20].

While preserving manuscripts is important, transcribing them is vital for the use and application of their contents. There have been several projects worldwide devoted to the transcription of ancient manuscripts into usable forms. In the mid 1600s, the Westminster Assembly met as a group of Presbyterian "Divines" claiming expertise in "divinity." They eventually wrote the documents upon which the Presbyterian faith was founded[21]. Manuscripts written by the members of the Westminster Assembly were found recently, including documents of correspondence, petitions, personal library lists, unpublished books, and early drafts of later published works. The Westminster Assembly Transcription Project was an effort to make the "unpublished manuscript writings of the

---

[19] Malik, Candra. 2009. Digital age provides hope for ancient manuscripts. *Jakarta Globe*, June 26, 2009. http://thejakartaglobe.com/news/digital-age-provides-hope-for-ancient-manuscripts/314733 (accessed September 11, 2009).

[20] Universitas Friburgensis. e-codices: Virtual manuscript library of switzerland. in Universitas Friburgensis [database online]. 2009 [cited 09/20 2009]. Available from http://www.e-codices.unifr.ch/en.

[21] Ryken, Phillip G., PhD. 1999. The westminster assembly (1643 - 1649).Tenth Presbyterian Church, Philadelphia, http://www.tenth.org/wowdir/wow1999-09-26.html (accessed October 13, 2009).

Westminster Assembly and its members as freely and widely available as possible[22]." Unfortunately, it is not a commercialized project yet as the founders have been asking for the help of interested and dedicated volunteers to aid in the full transcription of these documents. The volunteers are mailed copies of the documents, and are responsible for creating a word document transcription, which is then sent in electronic and paper form back to the project headquarters. All of the completed and verified transcriptions have become available for historians and the general public to access via the Westminster Assembly Digital Library[23].

Another effort known as the VENetIan Virtual Archive (VENIVA) was initiated in 1995 and lasted until 1998. Its goal was to preserve and improve the system of consultation of all historical documents in the Venice State Archive, in addition to the Marciana National Library of Venice, the Cephalonia County Archive, and the Austrian State Archive. This program aimed to create a web-based application that utilized remote access to make such documents readily available to the public. Additionally, various "cultural institutions," which includes libraries and other archives, would be able to access the application without charge or for a minimal fee. The final product resembles an interactive history lesson modeled from maps, letters, and drawings found in the participating archives. This project focused solely on digitizing the documents and using them to guide an individual through the history of Venice. It, however, did not delve into the transcription of the documents[24].

---

[22] Van Dixhoorn, C., Bower, J. 2008. *The westminister assembly transcription project.* The Westminister Assembly Digital Library, , http://www.westminsterassembly.org/digital-library/ (accessed 30 Sept 2009).

[23] Van Dixhoorn, C., Bower, J. 2008. *The westminister assembly transcription project.* The Westminister Assembly Digital Library, , http://www.westminsterassembly.org/digital-library/ (accessed 30 Sept 2009).

[24] VENIVA - the project. in VENIVA Consortium [database online]. 1996b [cited September 19 2009]. Available from http://www.tridente.it/progetto/home/project.html.

3.1.2 CURRENT MANUSCRIPT CONSULTATION PROCESS IN VENICE

Historians have centuries worth of preserved, handwritten documents available to them in Venice. All of the documents pertaining to the ancient city of Venice are housed in one central location: the Venice State Archive. It is located in the former convent of Santa Maria dei Frari, in Venice. It was established as the location for the archive between 1815 and 1822. Prior to the Archive being located at the Frari convent, it was located in the Ducal Palace. The individuals that organized the relocated archive at Frari instituted the same organization method that was present in the Ducal Palace[25]. There are approximately ninety (90) kilometers of shelves located in the archive, filled with vast amounts of documents, such as maps, letters, and governmental records from the ancient city[26]. The shelves are organized according to the office that created the documents and when they were created[27].

The wealth of information is astounding, and of great use to historians, who regularly visit the Archive to transcribe and study its contents. Documents that remain from the ancient city of Venice are an incredible asset to examining the history of Venice and its contemporary city-states. Records that span a total of eleven (11) centuries lie within the Venetian manuscripts, detailing the relations between these states. These accounts are considered to be among the most reliable sources of information on specific events and information regarding Venice. Additionally, a great deal of information is contained within the manuscripts that concern all of Italy and Europe. The sheer diversity of the types

FIGURE 9: STUDY ROOM AT VENICE STATE ARCHIVE

[25] Venice state archive. in Venice State Archive [database online]. 2007 [cited 09/30 2009]. Available from http://www.archiviodistatovenezia.it/index.php?id=46&L=3.

[26] Venice state archive. in VENIVA Consortium [database online]. 1996a [cited September 19 2009]. Available from http://www.tridente.it/progetto/services/venice.html.

[27] Venice state archive. in Venice State Archive [database online]. 2007 [cited 09/30 2009]. Available from http://www.archiviodistatovenezia.it/index.php?id=46&L=3.

of documents included in the Venetian manuscripts is astounding. Maps are just one example of such a document, which through analysis, can be seen to change over time in both style and technique[28].

While the Venice State Archive allows scholars and experts to examine documents from early time periods, the method is somewhat flawed. There is no efficient way to determine what information lies within a certain manuscript without manually searching page by page. Additionally, the manuscripts are not permitted to leave the study room in the Archive. This limits access for the historians that wish to utilize the information. They must be physically present in the Archive in order to access and study the manuscripts, creating an inconvenience during their research. Also, all of the information collected by a historian is strictly for his own personal use. There is no easy way to document his transcriptions for future. This means that many of the same documents are read and reread, transcribed and retranscribed, translated and retranslated[30].

FIGURE 10: COLLECTION OF DOCUMENTS AT ARCHIVE[29]

3.1.3 STATE-OF-THE-ART IN MANUSCRIPT CONSULTATION

To remedy each and every flaw of manuscript consultation in Venice, a computer program known as *uScript* was developed. It emerged under the Emergent Transcriptions Initiative at Worcester Polytechnic Institute as a senior project. The goal of the initiative was to "create a system that will leverage the work that is already being done by visitors to Archive around the world… [*uScript*] seeks to create an ever-growing collection of

[28] Italian National Memory of the World Committee. Memory of the world application: Italy - venice state archive. in Italian National Memory of the World Committee [database online]. [cited 6 October 2009]. Available from http://74.125.113.132/search?q=cache:RjR1yNpSuzEJ:portal.unesco.org/ci/en/files/17826/11038936681italy_venice.doc/italy_venice.doc+Memory+of+the+World+Programme+Venice&cd=1&hl=en&ct=clnk&gl=us.

[29] Image courtesy of http://www.archiviodistatovenezia.it

[30] Carlson, E., F. Nwaobasi, S. Saxena, and V. Truchanovicius. 2008. *Venice: Genesis of the city and its people.* Worcester Polytechnic Institute, (accessed September 15, 2009).

transcriptions that are refined by the work of many researchers[31]." The *uScript* program has three (3) main components, the Archive Assistant, the Transcription Assistant, and the Contribution Accountant, which work together to ensure qualitative searchable transcripts.

The various designers of *uScript* envisioned the program as an easy and effective way to search and transcribe the manuscripts of the Venice State Archive, or potentially any archive around the world. To begin, a user would use the Archive Assistant to query the server of the Archive, using already established metadata search standards. If the Archive has a page that matches these standards, then the server will respond to the query and a downloadable image will be made available to the user.

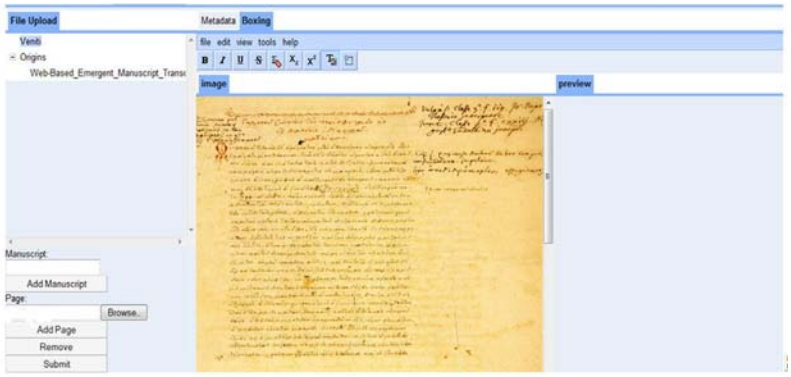


FIGURE 11: *USCRIPT'S* ARCHIVE ASSISTANT[32]

The Transcription Assistant feature of *uScript* is responsible for helping a user create a digital transcription of a particular manuscript. It uses adjustable boxes to frame words or phrases, which can then be transcribed by the user. The completed transcription would then be sent back to the *uScript* server, and linked to the original page for reference

[31] Cafferelli, B., Piper, N. and Sutman, E. uSCRIPT: Preserving history in a digital age. 2008 [cited 09/15 2009]. Available from http://www.uscript.org/newhome.html.

[32] Image courtesy of http://www.veniceprojectcenter.org

by subsequent users. Additional completed transcriptions allow users to search using both metadata standards and actual transcription keywords[33].

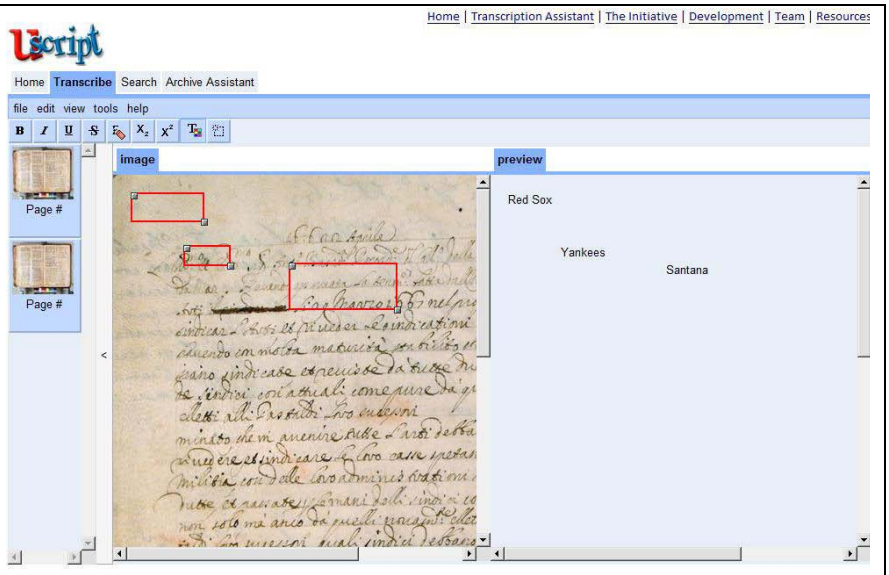


FIGURE 12: *USCRIPT'S* TRANSCRIPTION ASSISTANT[34]

To date, a great deal of work has been put into the coding of the *uScript* software. However, there is currently no working database for the Archive Assistant or concrete design for the Contribution Accountant. This last aspect of *uScript* is responsible for rating the transcriptions that are added to the database. These "reputation scores" would most likely be determined on the quality and quantity of transcriptions created by a user. The Contribution Accountant helps to add an aspect of accuracy to the transcriptions, so that users do not doubt the contents of the transcriptions. Additionally, there are no funds to allow for the completion of the software or to finance a support staff that would be responsible for the scanning of the documents and the general upkeep of the website[35].

---

[33] Ho, O., P. Chirag, R. Patel, and R. Klingman. 2003. *Manuscript transcription assistant.* Worcester, MA: Worcester Polytechnic Institute, Major Qualifying Project.

[34] Cafferelli, B., Piper, N. and Sutman, E. uSCRIPT: Preserving history in a digital age. pg 41. 2008 [cited 09/15 2009]. Available from http://www.uscript.org/newhome.html.

[35] Cafferelli, B., Piper, N. and Sutman, E. uSCRIPT: Preserving history in a digital age. 2008 [cited 09/15 2009]. Available from http://www.uscript.org/newhome.html.

3.2 METHODOLOGY AND RESULTS

uScript has the potential to change the way Venetians and historians examine the history of one of the most unique cities in the world: Venice. Several of its aspects need improvement, however, in order to create a fully functioning system. In order to obtain support for the *uScript* program, there needs to be some way for interested individuals to learn about the vision and the goals of the system. A promotional website was designed to display everything concerning *uScript*. The thought process for this website is discussed in section 3.2.1. Very little headway can be made on the coding and implementation of *uScript* without sufficient funds. The strategy employed for the completion of this objective is outlined in section 3.2.2. Through this continuing effort, it will be possible to turn *uScript* into a fully functional tool to aid in the transcription and preservation of the manuscripts of Venice.

3.2.1 REDESIGNED THE *USCRIPT* WEBSITE

This study aimed to promote *uScript* by designing a fully functional, informative and updatable website. Ultimately, scholars and experts could access the website to educate themselves on the efforts behind the *uScript* project in an effort to gain support for its implementation in the Venice State Archive. Previously, two (2) *uScript* websites were available on the internet, created by two (2) separate Worcester Polytechnic Institute project groups. It was determined that a combination of these websites would yield the ideal website. A Content Management System, CMS, was utilized to create a general *uScript* promotional website with improved overall appearance and performance. A CMS was ideal because it can be easily updated and modified. A WordPress website was created and a layout was selected for the *uScript* website, with the

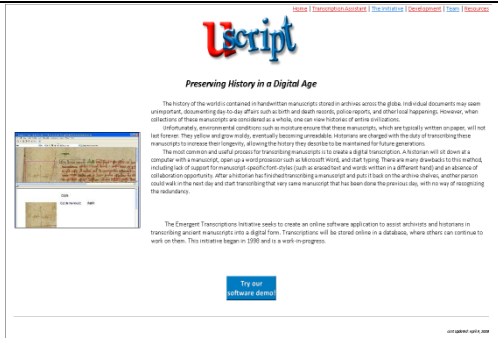


FIGURE 13: NON-CMS WEBSITE[36]

[36] Image courtesy of http://www.uscript.org

thought that it should be simple and functional, ultimately making the website more user-friendly.

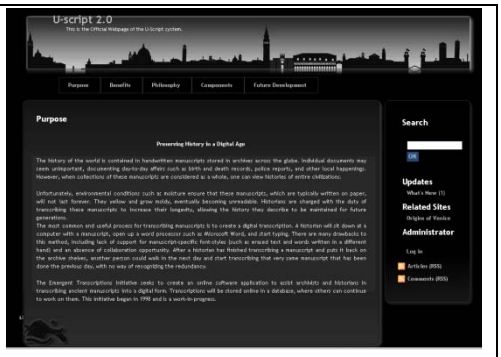


FIGURE 14: PREVIOUS CMS-BASED WEBSITE[37]

The promotional website became the main deliverable for the *uScript* aspect of this project. A great deal of time went into the improvement of the content of the site. The goal was to explain the development and the goals of the software, in both technical and promotional contexts. The main page is significantly more dynamic than the original sites, using flash animations to visually describe the transcription process. It integrates a blog, so that involved *uScript* users or future Worcester Polytechnic Institute project groups can create posts to showcase *uScript's* progress.

FIGURE 15: NEWLY REDESIGNED *USCRIPT* HOMEPAGE

[37] Image courtesy of http://www.veniceprojectcenter.org

A page is dedicated to each of the three components of *uScript*; the Transcription Assistant, Archive Assistant, and Contribution Accountant. The pages also feature flash animations that describe the components functionalities and the larger scope of the *uScript* application. There is also a new "Get Involved" section, which is designed to target experienced programmers that are willing to help develop various aspects of the *uScript* program.

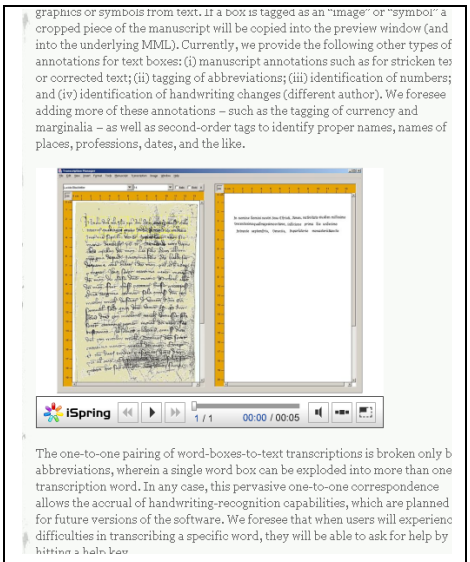


FIGURE 16: TRANSCRIPT ASSISTANT ANIMATION

New sections were added, titled Research, Funding and Resources. The Research tab has all of the papers written on *uScript*, in addition to links to web pages used to plan the development of the system. The Funding page has information on the current grant proposals, in addition to links to other grant resources. The Resources page has several links to pages of interest that are indirectly related to *uScript*, such as the Venice 2.0 web page and blog. These additional sections allow for all relevant information to be in a centralized location to increase the ease of access.

3.2.2 RESEARCHED FUNDING FOR THE DEVELOPMENT OF *USCRIPT*

Funding for computer programming and development will make *uScript* available for use in Venice. In the summer of 2009, a grant was submitted to the National Endowment for the Humanities to allocate funds for the development of *uScript*. A decision has not been made at this time, however, this study explored additional funding options. It would be in the best interests of the program to have as many public and private options as possible, in order to have the best chance of obtaining the necessary funds to launch *uScript*.

Through research, it was determined that the *uScript* proposal can be submitted to at least two (2) additional foundations, the Venetian Government and the Gladys Krieble Delmas Foundation. The Gladys Krieble Delmas Foundation deals solely with independent research on Venetian history and culture, offering up to $19,900 that could be used for the actual transcription of the manuscripts or for the development and implementation of the

open source coding[38]. The Venetian government, more specifically the office of archaeology for the Soprintendenza, offers money for projects for the improvement of preservation of the public domain. This includes everything from archaeology, public art, and libraries, therefore *uScript* would qualify. The grant itself would not have to be changed, only translated into Italian[39].

3.3 CONCLUSIONS AND RECOMMENDATIONS

While a great deal of progress was made with this project, there is still some work to be done. The promotional website accomplishes everything that it sets out to, from explaining the basics of the system to encouraging interested individuals to contribute to the program. What *uScript* really needs is coding for the design and the components to make it a functioning system, once and for all. At this point, there is a grant submitted to the National Endowment for the Humanities, which will provide funds to hire a programmer to work on the system. There should be a decision soon, and if it is a favorable decision, a programmer can get right to work creating the different features of *uScript*. If it is not favorable, the grant should be re-evaluated and submitted to other foundations, as previously suggested.

[38] Gladys Krieble Delmas Foundation. [database online]. 2008 [cited October 3 2009]. Available from http://www.delmas.org/

[39] Soprintendenza per i Beni Architettonici e Paesaggistici di Venezia e Laguna. [database online] 2009 [cited October 7 2009]. Available from http://www.soprintendenza.venezia.beniculturali.it/

4 GENETIC GENEALOGY IN VENICE

When written documents fail to explain the very beginnings of ancient civilizations, scholars look to scientists for answers. New advances in technology have given scientists the ability to trace the origins of populations through genetic genealogy. This technology is of great importance in understanding human roots. Today, countless genealogy websites exist for the sole purpose of tracing one's ancestry. Established organizations and companies have even merged to form super companies with the goal of tracing the ancestry of all of humankind. These super companies combine history and science to help shed light on the mysteries surrounding populations across the globe. The origins of Venetians in particular have long since been debated. Several theories exist that explain the migratory path of ancient populations in the Veneto but the historical references these theories are based upon are contradictory. This study aims to dispel the mysteries surrounding the origins of Venetians by using genetic genealogy.

4.1 BACKGROUND

Although several theories exist regarding the origins of Venetians, there is little historical evidence to confirm them. Genetic genealogy is the newest method to examine the oldest information: DNA. Genetic genealogy has made valuable contributions to world history in recent years by helping uncover the migratory path of ancient populations across the globe. By applying genetic genealogy process to Venetian DNA, scientists will be able to unravel the mysteries surrounding their origins. A background on genetic genealogy in Venice is presented in section 4.1, with the early history of the Veneto, Venetian origins theories, genetic genealogy processes, and the use of genetics to solve origins mysteries organized as subsections 4.1.1, 4.1.2, 4.1.3, and 4.1.4, respectively. The methods used and results obtained can be found in section 4.2. Lastly, conclusions and future recommendations for this project are presented in section 4.3.

4.1.1 EARLY HISTORY OF THE VENETO

FIGURE 17: THE VENETO REGION OF ITALY[40]

The history of Veneto is long and rich. The first documented people in Veneto were the Euganei, who populated central and northern Italy along the Alpine foothills, the Raeti, and Trentino and Alto Adige valleys[41] around 2,000 B.C.E[42]. According to legend, the Euganei were driven westward by a mysterious Adriatic population around 1,000 B.C.E. Around 700 B.C.E, this new Adriatic population merged with two additional peoples who inhabited the ancient Veneto; the Greeks and the Etruscans[43]. This merged population came to be known as the Veneti (now recognized as the paleo-Veneti), and soon, the Veneto became their established region.

Centuries later, the Veneto ceased to be a distinct territory of the Veneti and instead, became a territory of the Holy Roman Empire, when it was invaded and conquered by the Romans. This is when it is believed that the Veneto first adopted the Latin language and rich culture of Rome.

In the fifth century, Alaric the Goth and Attila and the Huns invaded Italy. Through their barbaric attacks on the city, Padua and Altinum were destroyed and burned to the ground. Many residents of Italian cities who were under heavy attack sought refuge in the Veneto lagoons, in provinces that would eventually become Grado and Venice. Torcello was

[40] Image courtesy of http://z.about.com

[41] *Euganei (people).* http://www.britannica.com/EBchecked/topic/195032/Euganei (accessed 11 10, 2009).

[42] *Euganei (people).* http://www.britannica.com/EBchecked/topic/195032/Euganei (accessed 11 10, 2009).

[43] Butterfield, Bruce J. *Livy's History of Rome.* 2006. http://mcadams.posc.mu.edu/txt/ah/Livy/Livy01.html (accessed 11 9, 2009).

established a safe-haven for refugees from the mainland in 569 A.C.E[44]. For this reason, Torcello is often called "the mother city of Venice."

After the Huns were defeated and driven out of Italy, many of the refugees returned to the Italian mainland to re-inhabit and rebuild their former cities. Then, in 586 A.C.E, Italy was invaded again, this time by the Germanic Lombards[46]. Italians again fled into the shallow waters of the lagoon where many remained, even after the Lombards were pushed out of Italy. These lagoon residents are credited as the founders of Venice, or more accurately as the first modern population to inhabit the Venetian lagoon.

FIGURE 18: ATTILA THE HUN'S INVASION OF ITALY[45]

But where did these lagoon people come from? And who was the Adriatic population that pushed the Euganei west and mixed with the Greeks and Etruscans in the Veneto? For centuries these questions have lingered without sufficient answers.

4.1.2 VENETIAN ORIGINS THEORIES

The origins of Venetians have been debated among historians for decades. Little evidence exists regarding the true origins of Venetians, and evidence that does exist often contradicts other historical references. Through research, however, three prominent theories have emerged that attempt to explain the origins of Venetians. They are the Paphlagonia Theory, Lusatia Theory, and Brittany Theory.

---

[44] Ammerman, Albert J. *Memoirs of the American Academy in Rome.* 2003. http://www.jstor.org/pss/4238806 (accessed 11 10, 2009).

[45] Image courtesy of http://thejournal.org

[46] *History of ITALY.* http://www.historyworld.net/wrldhis/PlainTextHistories.asp?gtrack=pthc&ParagraphID=glq (accessed 11 9, 2009).

PAPHLOGONIA THEORY

According to the Paphlagonia Theory, the Adriatic population that drove the Euganei out of the Veneto and merged with the Greeks and Etruscans was from Paphlagonia. Historical evidence suggests that following the great Trojan wars in the early centuries B.C.E, the Paphlagonians emigrated from modern-day northern Turkey, along the Adriatic and eventually into the Veneto. According to the Paphlagonia Theory, these peoples are the ancestors of modern-day Venetians.

The historical evidence that supports the Paphlagonia Theory is ancient. Both Livy and Homer provide literary references that support the Paphlagonia migration. As an ancient Roman historian, Livy kept historical records for much of Italy during the time of Augustus. According to Livy, a population called the Enetae emigrated from their homeland, in Paphlagonia, following the death of their King Pylaemenes in a revolution during the Trojan wars[47]. The Enetians, explained Livy, migrated westward along the Adriatic coast and eventually settled in the region that constitutes modern-day Veneto[48]. Livy claimed that once the Enetians arrived in the Veneto, they defeated the Euganei and established their own nation called Veneti. In *Ab Urbe Condita*, he writes:

FIGURE 19: PAPHLAGONIAN MIGRATION THEORY

[47] Butterfield, Bruce J. *Livy's History of Rome.* 2006. http://mcadams.posc.mu.edu/txt/ah/Livy/Livy01.html (accessed 11 9, 2009).

[48] (*Venetoimage - Canova e Possagno.* http://www.venetoimage.com/svc.htm (accessed 11 2, 2009).

FIGURE 20: PYLAEMENES[49]

"ANTENOR SAILED INTO THE FURTHEST PART OF THE ADRIATIC, ACCOMPANIED BY A NUMBER OF ENETIANS WHO HAD BEEN DRIVEN FROM PAPHLAGONIA BY A REVOLUTION, AND AFTER LOSING THEIR KING PYLAEMENES BEFORE TROY WERE LOOKING FOR A SETTLEMENT AND A LEADER. THE COMBINED FORCE OF ENETIANS AND TROJANS DEFEATED THE EUGANEI, WHO DWELT BETWEEN THE SEA AND THE ALPS AND OCCUPIED THEIR LAND. THE PLACE WHERE THEY DISEMBARKED WAS CALLED TROY, AND THE NAME WAS EXTENDED TO THE SURROUNDING DISTRICT; THE WHOLE NATION WERE CALLED VENETI[50]."

Homer also references the Enetae and Paphlagonians in the *Illiad*. In Book 2 of the *Illiad*, he writes:

"THE PAPHLAGONIANS WERE COMMANDED BY STOUT-HEARTED PYLAEMANES FROM ENETAE, WHERE THE MULES RUN WILD IN HERDS. THESE WERE THEY THAT HELD CYTORUS AND THE COUNTRY ROUND SESAMUS, WITH THE CITIES BY THE RIVER PARTHENIUS, CROMNA, AEGIALUS, AND LOFTY ERITHINI[51]."

In a similar manner to Livy, Homer explains that the Enetae were allied with the Trojans and aided them in the Trojan wars. He claims that the Enetae migrated westward from Paphlagonia following the loss of their king[52].

[49] Image courtesy of http://www.gutenberg.org

[50] Butterfield, Bruce J. *Livy's History of Rome.* 2006. http://mcadams.posc.mu.edu/txt/ah/Livy/Livy01.html (accessed 11 9, 2009).

[51] *The Classic Internet Archive | The Iliad by Homer.* http://classics.mit.edu/Homer/iliad.2.ii.html (accessed 11 15, 2009).

[52] *The Classic Internet Archive | The Iliad by Homer.* http://classics.mit.edu/Homer/iliad.2.ii.html (accessed 11 15, 2009).

Following both Livy and Homer's written evidence, the ancestors of modern-day Venetians are the Enetians, a Paphlagonian population that migrated to the Veneto after the loss of their king following the Trojan war.

LUSATIA THEORY

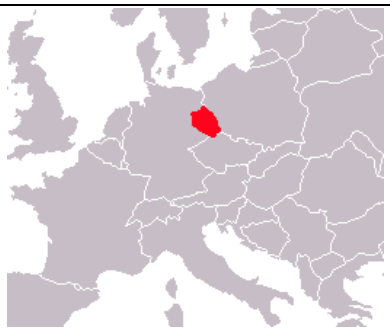


FIGURE 21: MODERN-DAY LUSATIA[53]

In contrast to the older and better documented Paphlagonia Theory, the Lusatia Theory is relatively new. It was first pioneered by Slovene author Jozko Savli. Through his research in the mid 1980s, Savli attempted to prove that Slovenes were not descendants of the Slavs, but in fact descendants of the Veneti. Savli's search came to be known as the Venetic-Theory. In 1984, he published *Veneti: First Builders of European Community* which explained his theory. His research was radical, as it evolved from just explaining the ancestors of Slovenes to explaining the ancestors of all of Europe. In his book, Savli explains that most current-day European populations are descendants of the Veneti, a central European population that emigrated from Lusatia across Europe to places such as Brittany, Austria, and Italy, among others. He attempts to associate each of the three distinct ethnic Veneti groups; the Adriatic Veneti, Vistula (Baltic) Veneti, and the Veneti of Gaul.

In short, Savli suggested that the Veneti were originally an ethic population located in the regions where modern-day Poland, Germany, and the Czech Republic intersect. The theory was highly criticized and rejected by many academics, but nevertheless, gave rise to much controversy and many questions.

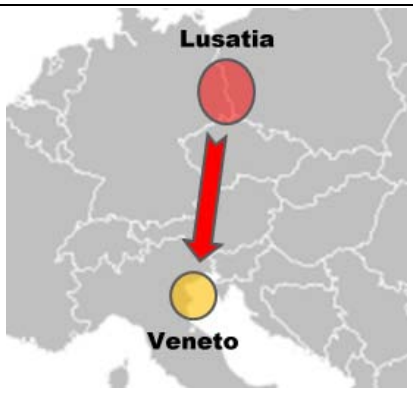


FIGURE 22: LUSATIA THEORY

Similarly, historian Stjepan Pantelic, argues that Slovenes were not descendants of the Slavs. His research stemmed from Savli's, in areas where Pantelic felt Savli's theory was

[53] Image courtesy of http://wikipedia.org

underdeveloped. Pantelic rejects the Paphlagonia Theory and instead claims that the Veneti came, not from eastern, but western Europe. According to Pantelic, "the Eastern Slavs are the most recent Slavs and there is no continuous record linking them to the Veneti mentioned at Troy.[54]" Following Pantelic's logic, the ancient Veneti only could have come from the basin of the Danube River. Pantelic suggests that the Veneti were indigenous people not of Paphlagonia, but of Carinthia, southern Austria. Pantelic further argues that the Veneti travelled to Paphlagonia to aid in the Trojan wars and returned home following the war, where they spread throughout western Europe and into the Adriatic.

Many academics reject Savli's theory, especially Roland Steinarcher, an Austrian-Italian historian. Steinarch claims that the Adriatic Veneti, Vistual Veneti, and Veneti of Gaul, are three distinct, unrelated populations[55]. He claims that the Enetae/Veneti translation was incorrect[56].

BRITTANY THEORY

The Brittany Theory is the least documented of each of the three Venetian origins theories. According to Strabo, an ancient Greek historian, modern-day Venetians are descendents not from Paphlagonia, nor Lusatia, but in fact, Brittany. Like Steinarcher, Strabo claimed that the Enetae/Veneti translation was incorrect and that the Eneti did not establish the nation of Veneti, as Livy claimed. Instead, Strabo suggest that the Veneti of Gaul, an ancient population of northeast France (modern-day Brittany) migrated to the

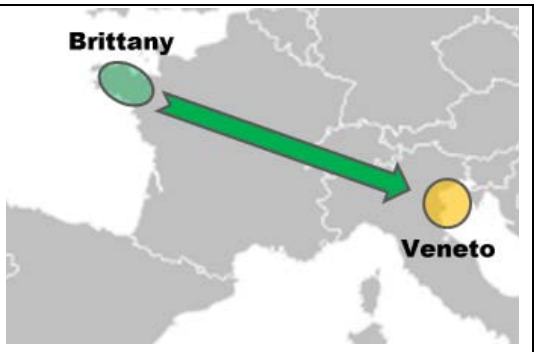


FIGURE 23: BRITTANY MIGRATION THEORY

[54] Pantelić, Stjepan. *Veneti Were Autochthonous in Carinthia.* http://www.korenine.si/zborniki/zbornik01/htm/pantelic.htm (accessed 11 2, 2009).

[55] Steinacher, Roland. *Studien zur vandalischen Geschichte.* June 2002. http://homepage.uibk.ac.at/~c61705/DISSERTATION-Volltext.pdf (accessed 11 3, 2009).

[56] Steinacher, Roland. *Studien zur vandalischen Geschichte.* June 2002. http://homepage.uibk.ac.at/~c61705/DISSERTATION-Volltext.pdf (accessed 11 3, 2009).

Veneto and are the ancestors of modern-day Venetians[57].

Contradicting evidence has left scholars at a standstill regarding the origins of Venetians. Since current historical evidence is based upon ancient records which are difficult to verify, scientific practices are being implemented to validate the theories.

4.1.3 CONTRIBUTIONS OF GENETIC GENEALOGY IN VENICE

Determining the origins of peoples using science is known as genetic genealogy, an application of genetics to traditional genealogy. In other words, genetic genealogy uses genetic testing to trace family lineages as far back as thousands of years[58]. Thus, genetic genealogy allows scientists to trace the migratory path of our common ancestors. More specifically, this new technology examines maternal and paternal lineages to determine the origins of our ancestors. In this way, researchers have begun to successfully trace the origins of modern populations to their ancient roots. Companies have even emerged with the goal of determining the origins of all of humankind through utilizing genetic genealogy. Through the collection of deoxyribonucleic acid (DNA), mysteries about modern populations can now finally be solved. Genetic genealogy is the most accurate means to solve the mysteries regarding the origins of Venetians.

The entire genetic genealogy process is long and complex. In short, scientists examine isolated fragments of DNA for small errors (called polymorphisms or mutations) in DNA. Since mutations are unique to different populations, they can be traced throughout history from generation to generation. It is through the comparison of mutations of hundreds of thousands of DNA samples that scientists can determine the origins of a population.

---

[57] *Strabo: The Geography.* July 18, 2009. http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Strabo/home.html (accessed November 18, 2009).

[58] Smolenyak, Megan Smolenyak. *Trace Your Roots with DNA.* United States of America: Rodale, Inc., 2004.

DEOXYRIBONUCLEIC ACID (DNA)

DNA is a long molecule that consists of basic units called nucleotides. These four nucleotides are adenine, guanine, thymine, and cytosine, abbreviated A,C,G, and T. DNA chains consist of millions of these nucleotide sequences arranged into two coiled strands that give DNA a double helix structure[60]. The nucleotides are complementary, meaning that adenine only pairs with thymine and guanine only pairs with cytosine. These pairings can be split apart when the cell wants to make new copies of DNA. The helix splits and the complementary nucleotides are generated via an enzyme called DNA polymerase, resulting in two (2) double helices from the original[61].

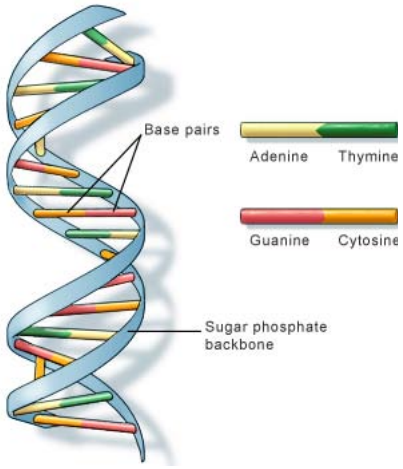


FIGURE 24: DNA STRUCTURE[59]

Stretches of DNA that contain important genetic information are called genes. The gene is defined as the unit of heredity[62]. Genes contain a specific sequence of nucleotides that generate a particular order of amino acids in a protein. Since proteins are essential to life, integrated into virtually every aspect of the human body, DNA is often considered the "blueprint of life[63]." Besides genes, DNA also contains long stretches of nucleotides that do not code anything. This is often called "junk" DNA[64]. As of 2004, nearly 95% of the human genome was identified as "junk" DNA, meaning that it does not have a purpose, or, more likely, that scientists have not identified its purpose[65]. The term "junk" DNA is deceiving, especially since it is within these non-coding stretches of DNA, called loci, that scientists examine for genetic genealogy.

[59] Image courtesy of http://writersforensicsblog.wordpress.com

[60] Smolenyak, Megan Smolenyak. *Trace Your Roots with DNA.* United States of America: Rodale, Inc., 2004.

[61] See footnote 58, page 53
[62] See footnote 58, page 53
[63] See footnote 58, page 53
[64] See footnote 58, page 53
[65] See footnote 58, page 53

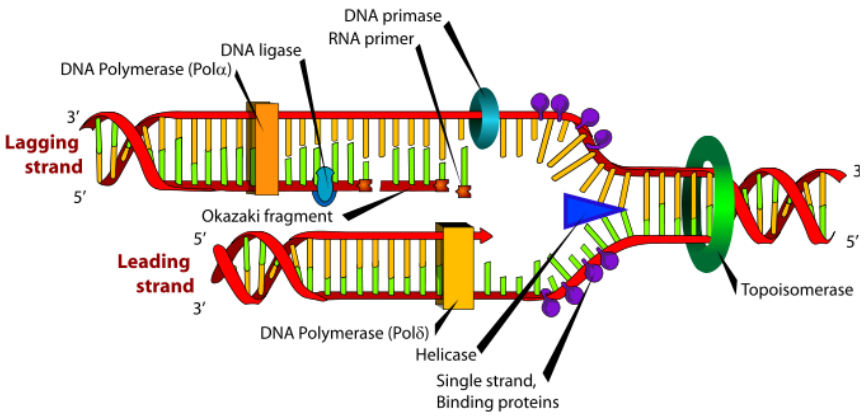


FIGURE 25: DNA REPLICATION PROCESS[66]

Y-CHROMOSOMAL DNA (Y-DNA)

Scientists are specifically interested in two types of DNA found within the human body; Y-chromosomal DNA and mitochondrial DNA[67]. It is through the analysis of these two types of DNA that scientists can determine the ancestry of humankind.

Chromosomes reside within the nucleus of a cell and contain human genetic information in the form of DNA. Humans consist of twenty-two (22) autosomes (called diploids) and one (1) sex pair of chromosomes. The female sex pair consists of two (2) X-chromosomes, denoted XX, while males consist of one (1) X-chromosome and one (1) Y-chromosome, denoted XY[68]. Because males only have one (1) Y-chromosome per cell, scientists often refer to them as haploids[69]. Since the Y-chromosome does not have a pair, it does not recombine, which means it remains virtually unchanged through the paternal lineage as the Y-chromosome is only present in males.

[66] Image courtesy of http://cellbiology.med.unsw.edu

[67] *The Genographic Project - Human Migration, Population Genetics, Maps, DNA.* 1996-2008. https://genographic.nationalgeographic.com/genographic/index.html (accessed October 10, 2009).

[68] *The Genographic Project - Human Migration, Population Genetics, Maps, DNA.* 1996-2008. https://genographic.nationalgeographic.com/genographic/index.html (accessed October 10, 2009).

[69] Sims, Jim. *What is Y-chromosome Analysis?* February 14, 2005. http://pages.sbcglobal.net/jimsims/Y-analysis/aboutYanalysis.htm (accessed 12 5, 2009).

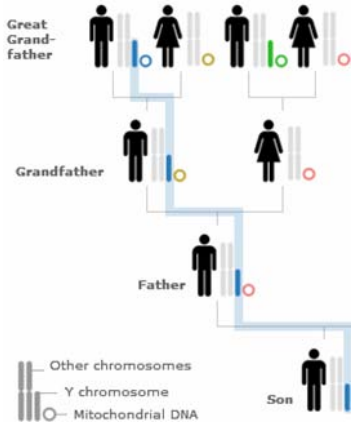


FIGURE 26: Y-DNA GENEALOGY[70]

In each pair, one chromosome is inherited from the father and one chromosome is inherited from the mother. These chromosomes recombine during meiosis, leading to generic variation in offspring[71]. This recombination of DNA makes the study of family trees difficult because the more the DNA recombines the greater the mixture and variance of genetic information from one generation to the next.

The process of copying DNA from one generation to the next is not perfect. Some minor mutations occur from time to time and it is these mutations that are eventually passed down from father to son in the Y-chromosome. The rate at which mutations occur is important to scientists. On average, mutations occur at a rate of once in every fifty million nucleotides[73]. Certain parts of Y-chromosomal DNA (Y-DNA) mutate at different frequencies, often above or below this average. If mutations occur too frequently, say once per 1,000 nucleotides, it would be difficult to compare Y-DNA since each Y-chromosome would be very different. On the other hand, if mutations occur too infrequently, say once per 100,000 years, it would be difficult to prove meaningful comparisons since almost all Y-DNA would be identical[74]. Scientists examine locations, called loci, along the Y-chromosome for mutations. A minimum of ten (10) loci are examined and compared in an effort to group Y-chromosomes based on their descent from a common ancestor[75].

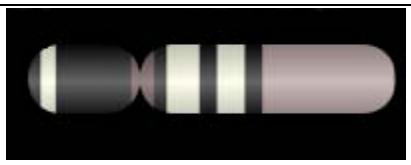


FIGURE 27: THE Y-CHROMOSOME[72]

[70] Image courtesy of http://genographic.nationalgeographic.com

[71] See footnote 69, page 55

[72] Image courtesy of http://wikipedia.org

[73] See footnote 69, page 55
[74] See footnote 69, page 55
[75] See footnote 69, page 55

Scientists inspect loci for sequences known as short tandem repeats (STRs). STRs are short patterns (two to five nucleotides in length) that are repeated in non-coding DNA[76]. For example, the four-nucleotide pattern **CATA** might be repeated three times, CATACATACATA. Occasionally, DNA polymerase, the enzyme responsible for duplicating DNA, loses its place and adds or subtracts a repetition[77]. Therefore, loci are examined for three specific STR occurrences:

1. Addition of a repeat
2. Deletion of a repeat
3. Addition or deletion of multiple repeats (rare)[78]

This copying glitch occurs, on average, two times out of one thousand. These STRs are what scientists examine to compare Y-DNA amongst males for genetic genealogy. Based on comparing just ten (10) loci, two (2) or fewer changes in genetic code (STPs) would indicate that the males are related. More than two (2) changes would indicate that males are not related. Examining more loci increases the statistical accuracy between samples[79].

Another less common mutation is called unique event polymorphisms (UEPs)[80]. These types of mutations occur so infrequently between generations (one in 50,000,000[81]) that scientists can treat UEPs as one-time events. In the Y-chromosome, these slow mutations occur in junk DNA. A particular UEP, called a single nucleotide polymorphism, or SNP, occurs when one single nucleotide is replaced with another, for example, a T for a G[82].

[76] See footnote 58, page 53
[77] See footnote 58, page 53
[78] See footnote 69, page 55
[79] See footnote 69, page 55
[80] See footnote 58, page 53
[81] See footnote 58, page 53
[82] See footnote 58, page 53

DNA analysis will output the names of each marker tested in loci in addition to how many STR repeats there are[84]. For example, a common DNA report might show DYS390: 14-12-24-11-13-12. DYS is an abbreviation for the particular STR that scientists are examining. The six most common markers, called classical markers, are DYS19, DYS388, DYS390, DYS391, DYS392, and DYS393. A specific combination of Y-chromosomal STRs is called a Y-DNA haplotype. A Y-DNA haplogroup, on the other hand, is a collection of people who share the same SNP[85]. Haplogroups are used to distinguish broad geographic categories, such as European or African. Haplotypes are less capable of distinguishing broad geographic categories since STR mutations are more common that SNPs, but help to narrow the scope of ancestry. Haplotypes and haplogroups are what researchers, like the ones involved in this study, are interested in when determining the origins of a population.

FIGURE 28: Y-DNA HAPLOGROUPS[83]

MITOCHONDRIAL DNA (MTDNA)

Similar to chromosomes, mitochondria contain genetic information in the form of mitochondrial DNA (mtDNA). Mitochondria are small energy producing structures found within the cytoplasm of a cell. MtDNA is considered the counterpart to Y-DNA since it allows scientists to examine the maternal lineage. These organelles replicate independently from the cell, and therefore, their DNA does not recombine during meiosis. MtDNA is the smallest of the DNA molecules, but its importance is far more significant than its size.

[83] Image courtesy of http://familytreedna.com

[84] See footnote 58, page 53
[85] See footnote 58, page 53

Hundreds, even thousands, of mitochondria can be found within a single cell in the human body[87]. Mitochondria are found in both males and females; however, offspring *only* inherit mitochondria from their mother. This phenomenon occurs during fertilization. Up to 100,000 mitochondria can be found in the female sex cell; the egg[88]. This is due to the nurturing role that the egg plays during fertilization. In contrast, the male sex cell, the sperm, only contains a few mitochondria. They are *only* present in its tail, called the flagellum[89], in order to provide energy to the sperm for transport. When the sperm penetrates the egg, the flagellum falls off. Therefore, male mitochondria are not passed to the zygote, which only receives mitochondria from the mother. MtDNA is thus passed down exclusively through the maternal lineage without recombination.

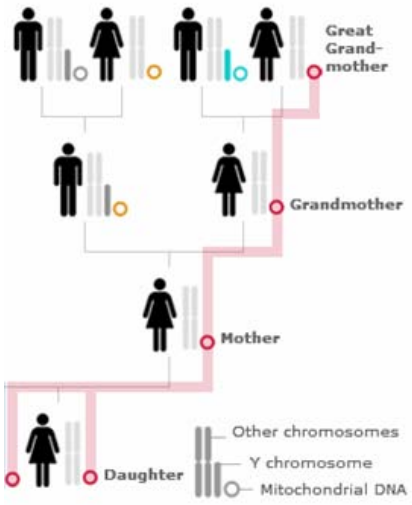


FIGURE 29: MTDNA GENEALOGY[86]

MtDNA forms a closed circle. Like Y-DNA, mtDNA consists of coding and non-coding sections. Scientists have dubbed the control region of mtDNA the Hypervariable Region, or HVR[91]. The region is named because of the high frequency of mutations that occur there. Scientists believe that the HVR acts like a spacer in mtDNA where the order of nucleotides does not seem to matter[92]. As most other regions of mtDNA are slow to mutate, the HVR is the most information in mitochondria.

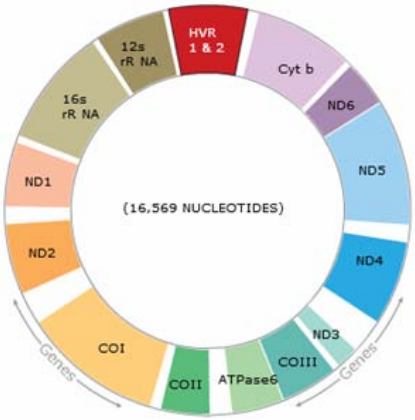


FIGURE 30: MTDNA STRUCTURE[90]

[86] Image courtesy of http://genographic.nationalgeographic.com

[87] See footnote 58, page 53
[88] See footnote 58, page 53
[89] See footnote 58, page 53
[90] Image courtesy of http://genographic.nationalgeographic.com

[91] See footnote 58, page 53
[92] See footnote 58, page 53

Scientists examine and sequence the entire HVR when examining mtDNA in humans. The results from the individual are then compared to the Cambridge Reference Sequence (CRS). The CRS was the first completely sequenced mtDNA, therefore it has been used as the standard for mtDNA comparision[94]. The complete set of mutations in the HVR is referred to as the mtDNA haplotype. Since mutations in the HVR are more common than elsewhere in the mtDNA, haplotypes do not provide the whole picture. MtDNA haplogroups, on the other hand, are the complete set of mutations found within the slow changing coding regions in mtDNA. MtDNA haplogroups, like Y-DNA haplogroups, distinguish broad geographic categories[95].

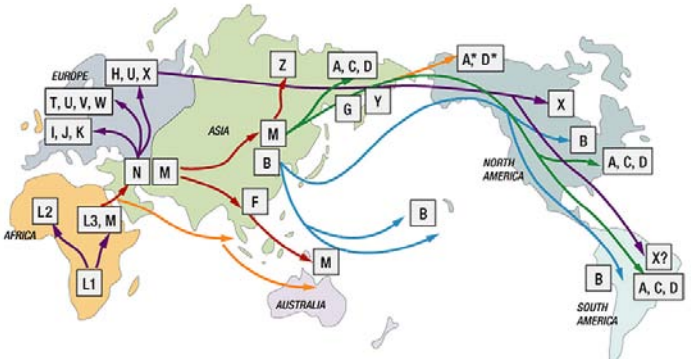


FIGURE 31: MTDNA HAPLOGROUPS[93]

4.1.4 USING GENETICS TO SOLVE ORIGINS MYSTERIES

In 2005, National Geographic and IBM launched the Genographic Project, a five year research partnership. The Genographic Project aims to "use cutting-edge genetic and computational technologies to analyze historical patterns in DNA from participants around the world to better understand our human genetic roots[96]." The project uses both Y-DNA and mtDNA analysis to trace the migratory path of humankind out of Africa. The project targets indigenous people from around the world for DNA sampling, but also invites the general public to submit their DNA anonymously through the purchase of a participation

[93] Image courtesy of http://genographic.nationalgeographic.com

[94] See footnote 58, page 53
[95] See footnote 58, page 53

[96] *The Genographic Project - Human Migration, Population Genetics, Maps, DNA.* 1996-2008. https://genographic.nationalgeographic.com/genographic/index.html (accessed October 10, 2009).

kit[97]. The participation kit includes a buccal swab, which is a non-invasive way to collect cheek epithelial cells to send to the Genographic lab for Y-DNA and mtDNA tesing.

Similarly, in 2000, National Geographic funded a genetic genealogy project that aimed to "learn whether modern Tyre, Lebanon fishermen are descendents from the ancient Phoenicians". To test their hypothesis, researchers collected DNA samples from modern fisherman in Lebanon, Tunisia, North Africa, and other locations where Phoenicians were believed to have visited and occupied. Scientists traced the mutations in the Y chromosomes of the collected DNA samples. Based on the genetic analysis, it was discovered that the Phoenicians were ancestors of today's Lebanese and occupied the region that stretches from modern-day Israel to Jordan. The research confirmed that modern Tyre fishermen are the descendants of the ancient Phoenicians. It is through advances in technology that research projects like the Genographic Project and the Phoenician Project have become possible.

Through collaboration with National Geographic and the Genographic Project, three hundred (300) DNA sample kits were given to a 2008 Worcester Polytechnic Institute project group for the collection of samples in northeast Italy, specifically the Veneto.

4.2 METHODOLOGY AND RESULTS

Contributing to the determination of the origins of the first inhabitants of Venice is a multi-step process. The first was to complete the Venetian DNA pilot study, for which the methods are described in section 4.2.1. The next step was to determine geographic locations, outside of Italy, where Venetians ancestors may still live today, which is shown in section 4.2.2. The final step, outlined in section 4.2.3, was to establish collaboration with the Genographic Project for the collection of future DNA samples from these geographic locations in an effort to determine the origins of Venetians.

[97] *The Genographic Project - Human Migration, Population Genetics, Maps, DNA.* 1996-2008. https://genographic.nationalgeographic.com/genographic/index.html (accessed October 10, 2009).

4.2.1 COMPLETED THE VENETIAN PILOT STUDY

Before the group could collect DNA samples from participants, a Worcester Polytechnic Institute Institution Review Board approval was required. An application was submitted prior to the departure to Venice and was approved. The approved WPI IRB form can be found in Appendix C.

Upon arrival to Venice, only sixty-three (63) of the DNA kits had been used. The group continued testing and administered one hundred (100) additional kits to Venetian males. The initial goal was to exhaust the remaining kits, however, this goal was modified after meeting with Genographic collaborator, Dr. David Comas, from the Universitat Pompeu Fabra, in Barcelona, Spain. According to Dr. Comas, one hundred sixty three (163) DNA samples would be sufficient for the Venetian pilot study analysis.

Samples were collected by targeting large Venetian organizations and through recommendations by the project advisor, Professor Fabio Carrera. ArcheoClub, Venessia.com, and the Settemari were three large Venetian organizations that were targeted for sampling. In addition, through the help of several individuals, the group was fortunate enough to obtain large quantities of DNA at various events and locations. These were obtained by attending Venessia.com's Funeral for Venice, visiting Caffé Brascilia, a Venetian café where many locals frequent. These organizations and events were targeted because they provided the best opportunity to collect male Ventian DNA. In addition to testing several large organizations, DNA samples were also collected from individuals based on recommendations of Professor Carrera. All of the individuals that were tested were given information regarding the prerequisites defined by the Genographic Project prior to the test, which can be found in Appendix D.

Once a test candidate was identified, the group administered the DNA sample kits. The testing protocol called for an informational form, consent form, questionnaire regarding ancestry of maternal and paternal grandparents, and a cheek swab. A detailed explanation of the testing protocol can be found in Appendix E. Participants were able to read the information sheet for addition details regarding the pilot study before signing their consent and filling out the ancestry questionnaire. Lastly, the group administered the

cheek swab and collected the saliva sample and cheek cell sample in a microcentrifuge tube.

To keep accurate records of test subjects, National Geographic included National Geographic Identification Codes (NGICs) which allow for the anonymous collection of DNA samples. One copy of the code was placed on the microcentrifuge tube, and a second copy was given to the participant on an information card which contained contact information for the Venice Project Center, Professor Carrera, and the Genographic Project website. An anonymous digital copy of the ancestry questionnaire was created for the Venice Project Center records. From this spreadsheet, the frequencies of ancestors from specific provinces within the Veneto can be identified. A confidential copy of the spreadsheet can be found in Appendix G.

In total, the group collected one hundred (100) DNA samples, bringing the grand total to one hundred sixty three (163). These samples were mailed to the Genographic Project Research Laboratory in Barcelona, Spain. In Barcelona, researchers and scientists will work to analyze and sequence various fragments of the DNA samples that the project group has collected. The information they gather will be added to the Genographic Project database. Individual participants who donated DNA samples for the Project will be notified by the Genographic Project with the details surrounding their ancestry using the NGIDCs.

4.2.2 IDENTIFIED FUTURE TESTING LOCATIONS TO VALIDATE ORIGINS THEORIES

The initial one hundred sixty three (163) DNA samples served as a pilot study and the base for potential DNA comparisons. Upon submitting the kits for processing, the group also included a map of future testing locations. These locations will serve as sites for future comparisons between the Veneto region and the possible origin regions, including Paphlagonia, Lusatia, and Brittany. After meeting with Dr. Comas, the group discovered that many samples from the Genographic Project are already available for DNA comparison amongst Paphlagonia, Lusatia, and Brittany, without the need of additional funding. Through extensive research, the group established a priority list of specific locations throughout Europe where Venetian ancestors are believed to have lived. The group researched ancient literature, as well as modern scientific papers on the theories of Venetians in an effort to identify references to Venetian ancestors. This produced a series

of maps of the different locations that the Venetians were thought to have migrated from. A screenshot of one of the maps can be found below in Figure 32:

FIGURE 32: THEORY REFERENCE MAP

4.2.3 ESTABLISHED FUTURE COLLABORATION WITH THE GENOGRAPHIC PROJECT

After discussing a future collaboration with Dr. Comas regarding the testing and comparison of DNA from regions that Venetians were believed to have come from, an agreement between the Venice Project Center and the Genographic Project was formalized. An agreement was signed by Professor Carrera and sent to Barcelona. Upon the completion of this study, the collaboration has not yet been finalized.

4.3 CONCLUSIONS AND RECOMMENDATIONS

The group was successful in completing the first objective, the Venetian pilot study. In total, over one hundred sixty three (163) DNA samples were collected. These samples were sent to Barcelona, Spain for DNA sequencing. Upon the completion of this project, the samples were not yet sequenced, so information regarding the origins of the Venetians could not be examined. The completion of the pilot study successfully ends the first step in using DNA to determine the origins of the Venetians. The next steps were to determine

which regions to compare Venetian DNA against as well as control regions, in order to accurately identify the origins of Venetians.

The group was also successful in identifying future testing locations to validate the origins theories, the second objective. The group created a custom Google Map, which can be found at *http://ve09orig.wordpress.com*. The origins theories maps can also be found at this website. In addition to the location and theories maps, the group developed a short journal article containing much of the evidence supporting each of the major theories. A copy of this journal article can be found at the aforementioned website. This unbiased journal article is one of the first to explore Venetian ancestry while showcasing each of the major theories based on scholarly research.

The last objective was to establish a future collaboration between the Genographic Project and the Venice Project Center. This collaboration was required for the comparison of Venetian DNA to Paphlagonia, Lusatia, and Brittany DNA. An agreement was reached between the Genographic Project and the Venice Project Center, but has yet to be finalized upon the completion of this study.

For the future, the group hopes that testing will continue, if needed, in the Paphlagonia, Lusatia, and Brittany regions. One hundred thirty seven (137) DNA sample kits remain from the initial three hundred (300) which could be used if future testing is required.

Upon receiving the results from these future test sites, the group recommends that a scientific journal article be published in a similar manner to the Phoenician Project, found in Appendix H. This journal article should include an explanation of each of the three origins theories, the methods used to collect DNA samples, and the scientific evidence that validates the true origin theory.

5 BIBLIOGRAPHY

Ammerman, Albert J. *Memoirs of the American Academy in Rome.* 2003. http://www.jstor.org/pss/4238806 (accessed 11.12.2009).

Archaeology. in Encyclopædia Britannica [database online]. 2009b (accessed 10.6.2009). Available from http://www.britannica.com/EBchecked/topic/32572/archaeology.

Bortoletto, Marco. Manuale di Procedura Archeologica. 2007, pg. 1.

Butterfield, Bruce J. *Livy's History of Rome.* 2006. http://mcadams.posc.mu.edu/txt/ah/Livy/Livy01.html (accessed 11.9.2009).

Cafferelli, B., Piper, N. and Sutman, E. uSCRIPT: Preserving history in a digital age. 2008 (accessed 9.15.2009). Available from http://www.uscript.org/newhome.html.

Carlson, E., F. Nwaobasi, S. Saxena, and V. Truchanovicius. 2008. *Venice: Genesis of the city and its people.* Worcester Polytechnic Institute, (accessed 9.15.2009).

Carver, M. 1996. On archaeological value. *Antiquity*(70): 45 (accessed 9.28.2009).

Djindjian, F. 1998. GIS usage in worldwide archaeology. *Archeologia e Calcolatori* (January): 19--30, http://soi.cnr.it/archcalc/indice/PDF9/09_03_Djindjian.pdf (accessed 9.28.2009).

Euganei (people). http://www.britannica.com/EBchecked/topic/195032/Euganei (accessed 11.10.2009).

Gladys Krieble Delmas Foundation. [database online]. 2008 (accessed 10.16.2009). Available from http://www.delmas.org/

History of ITALY. http://www.historyworld.net/wrldhis/PlainTextHistories.asp?gtrack=pthc&ParagraphID=glq (accessed 11.9.2009).

Ho, O., P. Chirag, R. Patel, and R. Klingman. 2003. *Manuscript transcription assistant.* Worcester, MA: Worcester Polytechnic Institute, Major Qualifying Project.

Italian National Memory of the World Committee. Memory of the world application: Italy - venice state archive. in Italian National Memory of the World Committee [database online]. (accessed 10.6.2009). Available from http://74.125.113.132/search?q=cache:RjR1yNpSuzEJ:portal.unesco.org/ci/en/files/17826/11038936681italy_venice.doc/italy_venice.doc+Memory+of+the+World+Programme+Venice&cd=1&hl=en&ct=clnk&gl=us

Malik, Candra. 2009. Digital age provides hope for ancient manuscripts. *Jakarta Globe*, June 26, 2009. http://thejakartaglobe.com/news/digital-age-provides-hope-for-ancient-manuscripts/314733 (accessed 9.11.2009).

Martinelli, Nicoletta, and Olivia Pignatelli. 1998. Datazione assoluta della piroga di lova (venezia). *Bollettino Del Museo Civico Di Storia Naturale Di Venezia* 49, : 207--212, http://www.archeosub.it/articoli/imbarcazioni/lovamonox.htm (accessed 9.18.2009).

Neal, C. 2009. *Beneath the corn, the husks of ancient venice*. The vigorous north. Vol. 2009, http://vigorousnorth.blogspot.com/2009/08/beneath-corn-husks-of-ancient-venice.html (accessed 10.2.2009).

Open context: Community-based data sharing and tagging. in Alexandria Archive Institute; UC Berkeley School of Information [database online]. b (accessed 9.18.2009). Available from http://opencontext.org/.

Pantelić, Stjepan. *Veneti Were Autochthonous in Carinthia.* http://www.korenine.si/zborniki/zbornik01/htm/pantelic.htm (accessed 11.2.2009).

Pulliero, Augusto. 2002. In laguna nord alla ricerca delle "radici". *Il Gazzetino Di Venezia*, http://www.archeosub.it/articoli/laguna/ammiana.htm (accessed 9.18.2009).

Rossi, Antonio. 1984. La laguna di venezia prima dei romani. *Archeologia Viva*(March), http://www.archeosub.it/articoli/laguna/lgnprst.htm (accessed 9.18.2009).

Ryken, Phillip G., PhD. 1999. The westminster assembly (1643 - 1649).Tenth Presbyterian Church, Philadelphia, http://www.tenth.org/wowdir/wow1999-09-26.html (accessed 10.13.2009).

Schloen, David. OCHRE: Online Cultural Heritage Research Environment. University of Chicago. (accessed 12.2.2009). Available from http://ochre.lib.uchicago.edu/index.htm

Sims, Jim. *What is Y-chromosome Analysis?* February 14, 2005. http://pages.sbcglobal.net/jimsims/Y-analysis/aboutYanalysis.htm (accessed 12.5.2009).

Smolenyak, Megan Smolenyak. *Trace Your Roots with DNA.* United States of America: Rodale, Inc., 2004.

Soprintendenza per i Beni Architettonici e Paesaggistici di Venezia e Laguna. [database online] 2009 (accessed 10.7.2009). Available from http://www.soprintendenza.venezia.beniculturali.it/

Steinacher, Roland. *Studien zur vandalischen Geschichte.* June 2002. http://homepage.uibk.ac.at/~c61705/DISSERTATION-Volltext.pdf (accessed 11.3.2009).

Strabo: The Geography. July 18, 2009. http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Strabo/home.html (accessed 11.18.2009).

Stratigraphy. in Encyclopædia Britannica [database online]. 2009d (accessed 11.10.2009). Available from http://www.britannica.com/EBchecked/topic/568372/stratigraphy.

The Classic Internet Archive | The Iliad by Homer. http://classics.mit.edu/Homer/iliad.2.ii.html (accessed 11.15.2009).

The Genographic Project - Human Migration, Population Genetics, Maps, DNA. 1996-2008. https://genographic.nationalgeographic.com/genographic/index.html (accessed 10.10.2009).

The Origins of Venice. Manuscript transcription. in The Venice Project Center [database online]. Venice, Italy, 2009 (accessed 11.10.2009). Available from http://origins.veniceprojectcenter.org/?page_id=39

Universitas Friburgensis. e-codices: Virtual manuscript library of switzerland. in Universitas Friburgensis [database online]. 2009 (accessed 9.20.2009). Available from http://www.e-codices.unifr.ch/en.

Van Dixhoorn, C., Bower, J. 2008. *The westminister assembly transcription project.* The Westminister Assembly Digital Library, , http://www.westminsterassembly.org/digital-library/ (accessed 9.30.2009).

Venetoimage - Canova e Possagno. http://www.venetoimage.com/svc.htm (accessed 11.2.2009).

Venice state archive. in Venice State Archive [database online]. 2007 (accessed 11.10.2009). Available from http://www.archiviodistatovenezia.it/index.php?id=46&L=3.

Venice state archive. in VENIVA Consortium [database online]. 1996a (accessed 9.19.2009). Available from http://www.tridente.it/progetto/services/venice.html.

VENIVA - the project. in VENIVA Consortium [database online]. 1996b (accessed 9.19.2009). Available from http://www.tridente.it/progetto/home/project.html.

White, Nancy. Types of archaeological data. 2004 (cited 9.28.2009). Available from http://www.indiana.edu/~arch/saa/matrix/ia/ia03_mod_05.html (accessed 10.3.2009)

GLOSSARY

Adriatic Veneti- an ancient population that inhabited northeastern Italy, in an area corresponding to the modern day Veneto region in Italy

Archaeological Potential- the likelihood that a particular area or region contains an archaeological artifact

Autonomous Agents- An autonomous agent is a system situated in, and part of, an environment, which senses that environment, and acts on it, over time, in pursuit of its own agenda. This agenda evolves from drives (or programmed goals). The agent acts to change the environment and influences what it senses at a later time. Non-biological examples include intelligent agents, autonomous robots, and various software agents, including artificial life agents, and many computer viruses.

Content Management System Web Page- a web page that utilizes content management system software for the creation and management of the HTML content; this type of software allows individuals with little or no knowledge of programming or markup languages to create and manage their own web pages easily.

Genealogy- the study of families and tracing of their lineages and history

Genetics- a discipline of biology, the science of heredity and variation in living organisms

Intervento/ interventii- excavation occurring within a site (i.e. Murano, Venice, etc.)

Meioses- process of reductional division in which the number of chromosomes per cell is cut in half, in animals it results in the formation of gametes.

Metadata- "data about data". Metadata articulates a context for objects of interest -- "resources" such as MP3 files, library books, or satellite images -- in the form of "resource descriptions".

National Geographic Genographic Project- launched on April 13 2005 by the National Geographic Society and IBM, is a five-year genetic anthropology study that aims to map historical human migration patterns by collecting and analyzing DNA samples from hundreds of thousands of people from around the world

Open Source- an approach to design, development and distribution offering practical accessibility to a product's source. In the case of *uScript*, this translate to open-source software, which is created and modified by developers who are willing to add functionalities to the system and build up on the freely provided code of the software.

Soprintendenza Archeologica per il Veneto, Nausicaa- entity in charge of archaeology.

Strato- layer

Veneti of Gaul- an ancient Celtic population best known for their battle with Julius Caesar, a coastal population that lived in southern Armorica (modern-day Brittany).

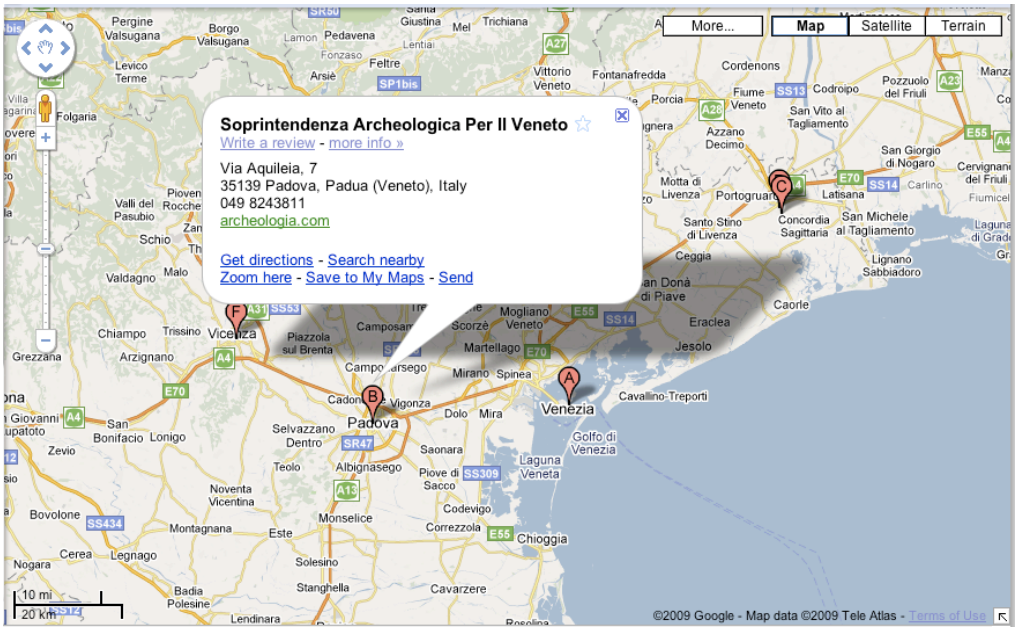
Vistula (Baltic) Veneti- an ancient population identified as inhabitants of modern-day Poland, along the Vistula River. Pliny the Elder, Ptolemy, the Polish archaeologist Okulicz, and the Gothic author Jordanes, all reference Veneti in the upper Vistula region in Poland. The Vistula Veneti are recognized as an ancient Indo-European people.

WPI- Worcester Polytechnic Institute

Zygote- the initial cell formed when a new organism is produced by means of sexual reproduction, the union of two gametes that fertilize and produce a diploid cell

Soprintendenza per i Beni Archeologici per il Veneto-

Via Aquileia 7 - 35139 PADOVA
tel. 049/8243811 - fax 049/8754647
archeopd@arti.beniculturali.it



ARCHEASY GRANT PROPOSAL

FIRST DRAFT

ARCHAEOLOGY MADE EASY

Worcester Polytechnic Institute

Venice Project Center

Principal Investigator: Professor Fabio Carrera

TABLE OF CONTENTS

Table of Contents IV
Project Description- ArchEasy V
Narrative V
Significance VI
Significance to Archaeologists VII
Significance to the Soprintendenza Archeologica per il Veneto VII
Significance to Public Works VIII
Significance to Archaeological Cross-Referencing and the Gain of Historical Knowledge VIII
Scope and Duration VIII
Methodology and Standards IX
System Components XI
Forms XI
Personalized Workspace XIII
Database XIII
Modeling Application XV
Staff XVII
Previous Grant History XVII

PROJECT DESCRIPTION- ARCHEASY

This proposal seeks to develop a system that centralizes archaeological data in a web-based application, known as ArchEasy. Our goal will be accomplished by implementing already existing technologies to create an archaeological database, advanced search functionalities, and user-friendly modeling tools to enable archaeologists to extract the most history from an intervento, or archaeological excavation. Various systems exist across the world for the management of archaeological data, but they do not offer all the features of logging, searching, and modeling such data. Additionally, each nation has its own laws regarding the collection of archaeological data, and thus each system must be relatively specific to a country. We propose ArchEasy to contain all these features, to be initially used by Venetian archaeologists, and eventually to allow users to modify it so it may be adapted to country.

Currently in Venice, the presence of an archaeologist is required at all construction sites. Artifacts are often found inadvertently during public works projects, thus stressing the importance of having someone there to record the data. Venetian archaeologists must follow the administrative guidelines set forth by the Soprintendenza Archeological per il Veneto Nausicaa. Various forms must be filled out, beginning with the Scheda di Intervento Archeologico. The Soprintendenza also mandates that a Unita Stratigrafica form (US) is filled out for every layer that is unearthed, requiring numerous details and photographs. For every object and layer recorded by the archaeologist, an expert from the Soprintendenza then completes an RA form. This system is very complex and the data is not easily accessible, thus making it difficult for archaeologist to cross-reference data.

Using ArchEasy, archaeologists will fill out all the required forms online. The information they enter will automatically be stored in the system database. This system will also accurately visualize recorded interventii, or archaeological projects, on a map of Venice. Along with this feature, the system will have a sophisticated search function that gives archaeologists the ability to easily access their reports. The system will also be able to analyze archaeological data to do such things as to increase the accuracy of calculating archaeological potential, and to group facts by different characteristics to recognize commonalities.

An Interactive Qualifying Project group at Worcester Polytechnic Institute (WPI) began ArchEasy in 2008. Currently, the system exists as a prototype website demonstrating its future functionalities. We eventually plan to make ArchEasy available to archaeologists, the Soprintendenza Archeologica, and the general public. Work will take place over a three-year period, with year one spent laying the groundwork and establishing software development frameworks, year two spent programming the key elements of the system, and year three testing, deploying and expanding the advanced capabilities of ArchEasy.

NARRATIVE

ArchEasy: A web-management system for the archival and visualization of archaeological data

To gain insight into modern day humanity, we must first examine the roots and traditions that define culture and people. Archaeologists devote their lives to examining ancient architecture and relics in an effort to understand long-ago civilizations. Venice in particular is a city like no other, where the rich culture and vibrant traditions of the past fuse together with the hustle and bustle of the modern world. All public construction requires the presence of an archaeologist in an effort to unearth and document potentially significant artifacts. Archaeological experts have attempted to create centralized archaeological potential maps that track the locations of such ancient artifacts found beneath the modern city, but these are often outdated. Cross-referencing and dating is made difficult without a system to store and organize all of the collected archaeological data.

Our project aims to improve the archival and visualization of archaeology in Venice by programming ArchEasy to most effectively aid archaeologists in storing and organizing data. The project we are proposing is predicated upon the following assumptions:

- A breadth of archaeological artifacts are discovered in the Veneto region
- The archaeological analysis and data collection of every intervento is then submitted to the Soprintendenza Archeological per il Veneto
- This information is not easily accessible by experts in the field due to its sheer abundance and lack of a search system
- Archaeological potential maps that show the location of artifacts are not easily updatable

Our solution to these problems is to put in place a system that will make the organization and search of archaeological data simple and effective, so that archaeologists can easily create a personalized workspace and also cross-reference other finds. An archaeologist inputting data into the system and creating interventii maps will have the option to allow other users to view this information. Data will thus be shared among archaeologists, and connections between finds will become apparent. The database will be constantly changing to reflect the current knowledge and finds in the archaeological field. By linking a modeling application to ArchEasy, an interactive archaeological potential map can be created to show the latest information. This will benefit public works officials and companies who wish to assay a site for archaeological value before initiating a project. An autonomous agent approach will be developed in which each object and layer is capable of reasoning for itself. In this way, objects correlate with one another without the need of a greater entity to direct them. This will in essence achieve a level of automatic referencing that is unparalleled in other archaeological web management systems.

SIGNIFICANCE

Our system aims to facilitate the archiving and organization of archaeological data in the city of Venice, Italy. To date, Venetian archaeologists fill out forms by hand, and must manually submit them to the Soprintendenza Archeologica per il Veneto. The data contained in these

forms is then difficult to locate and access by other experts in the field. Since understanding the history of a relic often relies on cross-dating techniques, we propose to implement a web management system to not only allow archaeologists to organize their finds, but to also search through the ever-expanding database. We want to give archaeologists a personalized workspace in which to complete administrative requirements as well as to interpret finds, to ultimately revitalize archaeological processes and shed light on Venice's history.

SIGNIFICANCE TO ARCHAEOLOGISTS

Thanks to ArchEasy, archaeologists will have an online space in which to manage and interpret their finds. In such a historically rich city as Venice, archaeological finds are frequent and plentiful. Given the large amount of administrative requirements dictated by the Soprintendenza Archeologica per il Veneto, documentation of an archaeological find is very arduous. Additionally, archaeologists must document their data multiple times, for their own records and also for the official records.

The system we propose will allow archaeologist to fill out the required documentation on one easy-to-use website. This information will automatically be entered into the ArchEasy database, so that there will be no need to reenter any data. The ability to create an interactive map for interventii will provide archaeologists a visual way to search through their data. They will also have the ability to upload GIS layers and objects, to be integrated into the creation of an archaeological potential map.

SIGNIFICANCE TO THE SOPRINTENDENZA ARCHEOLOGICA PER IL VENETO

Currently in Venice, the Soprintendenza Archeologica oversees the archaeological projects in the city. The Soprintendenza Archeologica will have an administrative user role, enabling it to access all of the completed interventii in the ArchEasy database. This administrator will also be in charge of managing users, by editing their information and adding and deleting users. As ArchEasy develops and gains popularity, it will act as an intermediary between archaeologists and the Soprintendenza. Necessary paperwork will become available to the Soprintendenza as it is completed, thus speeding up the administrative process. These records will be electronic, thus easily organized, stored, and subsequently searched. Our system provides archaeologists with a space to organize and visualize their finds, to present them in this way without added effort to the Soprintendenza Archeologica.

SIGNIFICANCE TO PUBLIC WORKS

Our system will provide public works officials and companies with an updated archaeological potential map, showing the relative locations of buried archaeological artifacts. This will greatly aid in the planning of construction projects, saving these entities both time and money. If a building planner were to determine the best location for initiating construction, he would first consult the ArchEasy map. It would alert him to any previous and current interventii in that region. If there have been many in one location, the planner may choose another region to build in order to avoid the danger of encountering another site of archaeological significance. This feature of ArchEasy will grant easy access to otherwise elusive information to individuals outside of the archaeological field.

SIGNIFICANCE TO ARCHAEOLOGICAL CROSS-REFERENCING AND THE GAIN OF HISTORICAL KNOWLEDGE

Individual artifacts may seem relatively unimportant, but when these are analyzed collectively, they can provide clear and insightful perspective on the history of an entire civilization. ArchEasy will provide a communal area for archaeological data to be deposited and interpreted. Dating in Venice, as well as other parts of the world, is often accomplished by cross-referencing other objects and by placing finds in historical context. If two walls are discovered that cross but do not weave, then it can be determined that these walls are from different time periods. If these walls are in fact woven, then the construction details must be studied to determine if they are of the same or different epochs. This includes the size of the bricks, the material, and the mortar itself. Similarly, in dating ceramic fragments, the fill is often used to determine the age in which it was produced. It is known with certainty that certain fragments are made only during particular time periods, thus narrowing the possibilities of their age. Historical events give power to the relative dating system. The more facts and objects correlate, the easier it becomes to accurately date them. Herein lies power of ArchEasy in bringing this information together and providing archaeologists with the tools to search though it. By adapting an autonomous agent approach to archaeological finds, ArchEasy will eventually suggest relationships to the archaeologist. This new technology is not intended to replace the judgment of the experts, for archaeology is an interpretative field that requires careful consideration. However, it will act as a tool to proposing possible relationships, to then be more carefully examined by archaeologists. This innovative way of interpreting data highlights the importance of developing its technologies.

SCOPE AND DURATION

We are requesting a three-year research grant to make our system fully operational so that it can begin to fulfill on its promises. The final product will greatly accelerate the production of archaeological results, by increasing the efficiency of the archival, geo-referencing, and interpretation of archaeological data. Over this three-year project, we foresee additional academic research projects will take place to explore the best technical alternatives ArchEasy's key features. While the needed research is being conducted, the bulk of the grant will be dedicated to developing the database, the sophisticated search mechanism, the archaeological potential map software, the autonomous agent technology adapted to archaeology, and the best user interface.

METHODOLOGY AND STANDARDS

Our ArchEasy project seeks to create an online application to create a workspace for archaeologist in which to effectively organize their finds and to increase the ease of relative dating by the cross referencing of data. The most advanced technologies will be used to create a system allowing archaeologist to fulfill administrative country requirements while creating a searchable database that integrates various types of data to ultimately reveal new relationships between their finds. Following are sections detailing some of our development philosophies and principles.

Current Standards in Archaeological Data Management

When data is made available through one predominant location, it can be collectively analyzed and modeled. Tools such as GIS, or Geographic Information Systems, have become extremely useful to the archaeological process. Archaeologists mostly use GIS and graphics for rescue archaeology, large site excavations, regional settlements studies and Cultural Resources Management (C.R.M.)[98]. Different GIS applications are available for the varying needs of its users, with the main goal of managing various graphic objects. Archaeologists can build a GIS site that allows direct interaction with the excavation data, management of this data, and also provides 3D modeling tools.

[98] (Djindjian, F. 1998. GIS usage in worldwide archaeology. *Archeologia e Calcolatori* (January): 19--30, http://soi.cnr.it/archcalc/indice/PDF9/09_03_Djindjian.pdf (accessed September 28, 2009).

Various archaeological web management systems are available from different areas of the world. Open Context[99] is one such system that allows its users to input primary field research data from archaeology and related fields. By making this data easily available electronically, it can be found and reused in the future by others without any hassle. Open Context hosts many different projects, with different methodologies, recording systems, and temporal and geographic coverage. It also uses many web-services so that data can be transferred back and forth between Open Context and other applications. Perhaps a very unique aspect of this service is the way in which it facilitates information retrieval by the use of Archaeological Markup Language (ArchaeoML). This system uses an item-based information model, where individual atomic units of observation are related to each other and their descriptions. This allows for deeper searching, with a variety of linked relationships displayed, not only the very obvious. ArhcaeoML is one of the technologies that may be adapted to ArchEasy.

The Online Cultural Heritage Research Environment" (OCHRE) is another such system, specifically an Internet database system for cultural heritage information started at the University of Chicago by David Schloen. The most unique feature of this system is its ability to integrate data of diverse origins by organizing into a coherent and searchable body of information[100]. The administrator of a particular project can import data tables into the OCHRE system and also export data from the central database at any time. OCHRE uses the XML Query Language to search and manipulate the data stored in the central database, which is preferable when working with hierarchical semi-structured data used in cultural heritage research, such as archaeology. Each XML document in the database corresponds to a real-world entity such as an archaeological site, context, or artifact, etc. These items are grouped into several general categories such as "Locations & Objects", in order to provide the most efficient storage and retrieval of the information. Data is linked in this way, and can be integrated from a single project to many different ones. The item-based data integration employed by the OCHRE system could be adapted to ArchEasy to allow archaeologists to store and use their data most effectively.

Hierarchy of users

User Type	Description

[99] Open context: Community-based data sharing and tagging. in Alexandria Archive Institute; UC Berkeley School of Information [database online]. b [cited September 18, 2009 2009]. Available from http://opencontext.org/.

[3] Schloen, David. OCHRE: Online Cultural Heritage Research Environment. University of Chicago. Accessed December 2, 2009. Available from http://ochre.lib.uchicago.edu/index.htm

Technical Administrator	• Access to back end of the ArchEasy system • Access to all database and user information • Receives requests for new users • Special URL to login
Soprintendenza Administrator	• Access to all database and user information • Special URL to login
Archaeologist	• Access to their own interventii and other publicly available interventii • Ability to set public permissions on personal data • Log in directly on ArchEasy website
Guest	• Only access to data made available to public by registered users • Must login as Guest directly on ArchEasy website

TABLE 2: HIERARCHY OF USERS AND DESCRIPTIONS

SYSTEM COMPONENTS

The system we are proposing consists of the following main features:

- **Forms** that are filled out electronically and easily saved.
- **Personalized workspace** that allows archaeologist to create interventii and set permissions.
- **Database** generated from information extracted from forms, searchable by users.
- **Modeling application** that allows archaeologists to visually display their interventii and highlights relationships between objects and layers

FORMS

This section of ArchEasy can be accessed from the Create Intervento menu tab. All of the administrative forms required of Venetian archaeologists by the Soprintendenza Archeologica per il Veneto can be found here, in their electronic versions. These can be filled out on the website, and the information subsequently stored in ArchEasy's database. If desired, indexes will automatically be created from these forms in order to organize all the documents. Other forms may also be found here, and eventually, there will be an option to allow archaeologist to

create their own forms. In this way, ArchEasy will become useful to experts from all over the world.

The following forms are currently in the system:

Form	Description
Scheda Sito	General information form, with five sections
Relazione	Relationships
Indice Foto	Index form of all the photographs
Foto Digitali	Digital photographs taken during the excavation
Foto Storiche	Historical photographs
Indice US	Index form of all the completed US forms
Schede US	Unita stratigrafica, mandated by Soprintendenza Archeologica to be filled out for every object and layer
Scheda Tomba	To be filled out for any human remains
Organico	To be filled out for any organic finds
Inorganico	To be filled out for inorganic finds
Statue Form	To be filled out for any statue or statue remnants
Mappe Storiche	Images of historical maps
Disegni	Design
Analisi	Summary of various kinds of analysis (chemical, dendrochronological, pedological, geological, anthropological)

TABLE 3: FORMS IN THE ARCHEASY SYSTEM

PERSONALIZED WORKSPACE

Archaeologists registered on ArchEasy will have at their fingertips the tools to create new interventii as well as to edit previous ones. Once the archaeologist has named the new intervento and entered its basic information into the system, such as director and site, he or she is free to fill out whichever forms are necessary for that excavation as many times as necessary. The completed forms will automatically be linked to that intervento, and can subsequently be viewed when the intervneto is accessed. Additionally, the archaeologist user can upload GIS images, drawings, and photographs related to the intervento. He or she can then set permissions for each of these items, to make them publicly available or to restrict their access according to his or her wishes. As ArchEasy becomes more developed, an archaeologist's userspace will become more and more customizable, to give everyone the necessary freedom to maneuver the breadth of data.

DATABASE

The ArchEasy database will be automatically generated from the information entered by users in creating interventii. This information will be extracted from the administrative forms, and from the files uploaded into the system. In the development of ArchEasy, the best standards will be chosen and applied to allow for the most diverse data to be effectively integrated. Consequently, this data can be accessed by user queries in the search feature of the system, as can be seen in Figure 1.

FIGURE 33: RUDIMENTARY EXAMPLE OF ADVANCED SEARCH FUNCTION

By adopting already-existing search standards, ArchEasy strives to yield search results that are based on many levels of data connectivity. In the demonstration website example, if Romano era is searched, then the results seen in Figure 2 would be obtained. This database will allow experts to access each other's data in order to better interpret their own finds. Our hope is to display relationships and search results that are not expected, but relevant, in order to improve the cross-dating process so prevalent in archaeology.

Select	Intervento	Century	Origins	Depth	Era
☐	Lazaretto Nuovo Intervento	XV	Rome	36 cm	Romano
☐	Murano Intervento	XIV	Rome	130 cm	Romano
☐	San Marco Intervento	XIII	Rome	39 cm	Romano
☐	Arsenale Intervento	XII	Rome	47 cm	Romano
☐	Piazza La Roma Intervento	XI	Rome	12 cm	Romano
☐	Lido Intervento	X	Rome	81 cm	Romano
☐	San Zaccharia Intervento	IX	Rome	10 cm	Romano
☐	Ferrovia Intervento	VIII	Rome	65 cm	Romano
☐	San Pietro Intervento	VII	Rome	100 cm	Romano
☐	Santa Margarita Intervento	VI	Rome	36 cm	Romano

Export Data into List | Map Results

FIGURE 34: SCREENSHOT OF EXAMPLE SEARCH RESULTS

MODELING APPLICATION

The interactive archaeological potential map is an important feature of ArchEasy that provides users with a visual representation of the interventii in a particular site. This live map will contain placeholders for interventii, as in the Google Map application. Once an intervento has been created and given a precise location, it will appear on the ArchEasy archaeological potential map as a new placeholder. When this icon is clicked, a window will pop up, displaying the particulars of the excavation and drop-down menus, enabling the user to view all associated items, as can be seen in Figure 3.

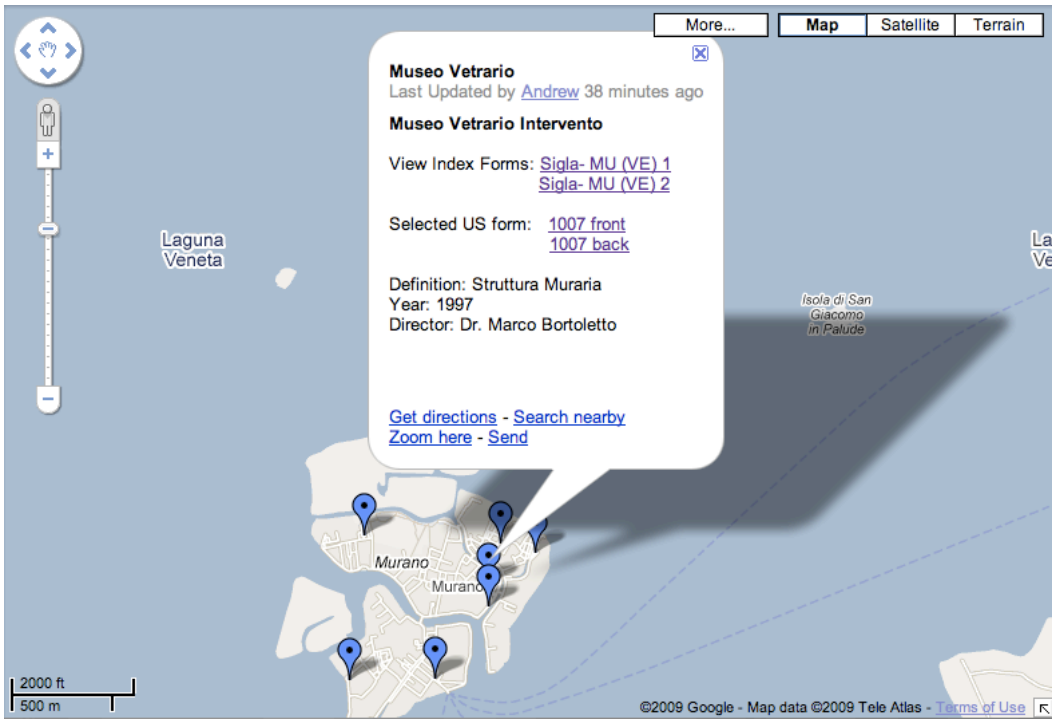


FIGURE 35: INTERVENTO MAP EXAMPLE SCREENSHOT

Eventually, there will be a GIS map of every site available through the ArchEasy system upon which archaeologist can overlay GIS representations of their interventii and all associated layers and objects. Autonomous agents will then be employed to enable these layers and objects to "think" and reason their relationships with one another.

STAFF

Fabio Carrera, PI

Stephen Guerin, consultant

Marco Bortoletto, consultant

PREVIOUS GRANT HISTORY

ArchEasy has not received any funding in the form of grants from any federal or nonfederal sources. Minimal funding (amounting to less than $1,000 over the past ten years) has been received from Worcester Polytechnic Institute in the form of student project fees.

APPENDIX B: SCREEN SHOTS OF ARCHEASY WEBSITE

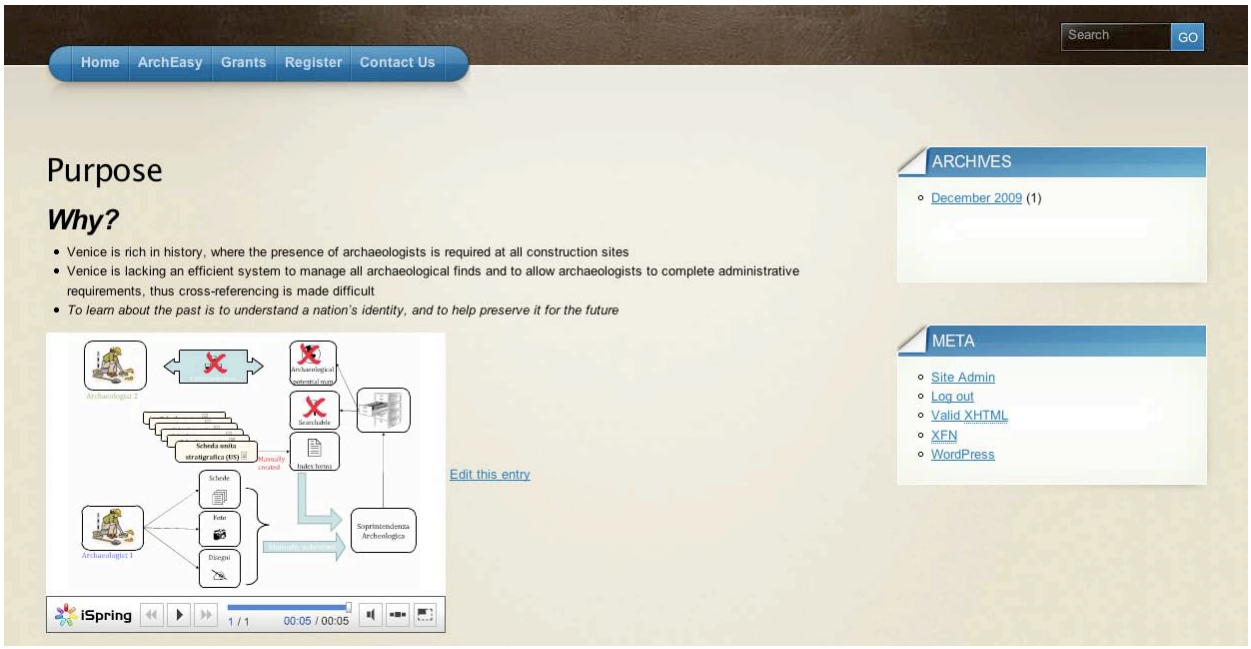


FIGURE 36: ARCHEASY PURPOSE

FIGURE 37: ARCHEASY VISION

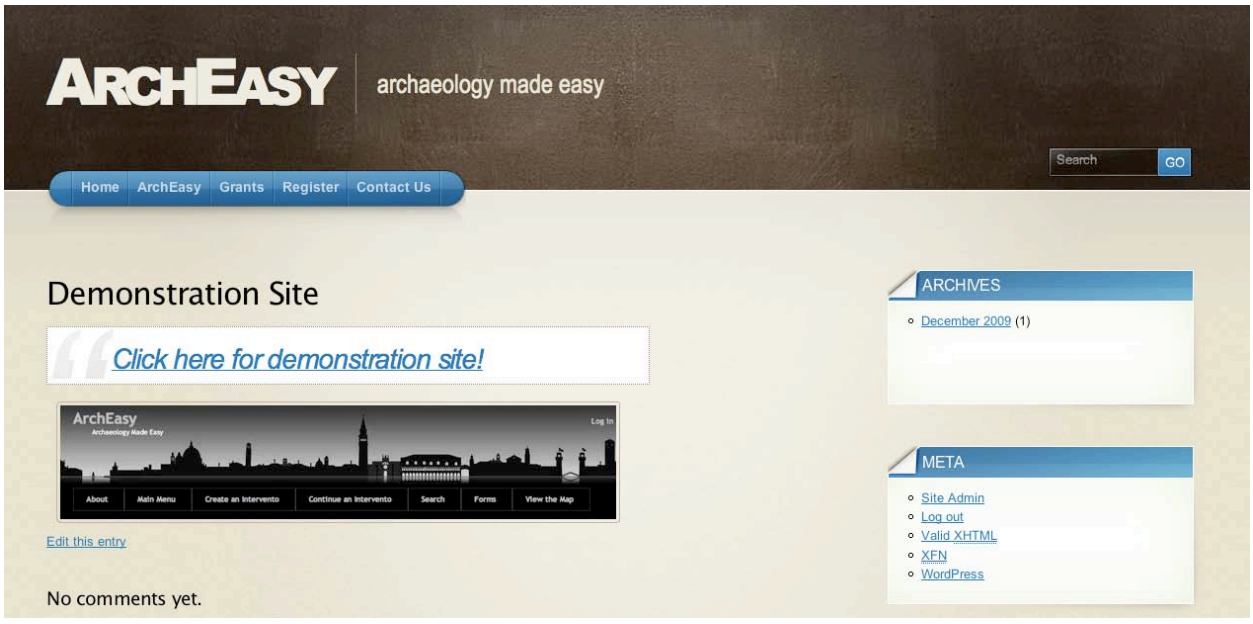


FIGURE 38: ARCHEASY DEMONSTRATION

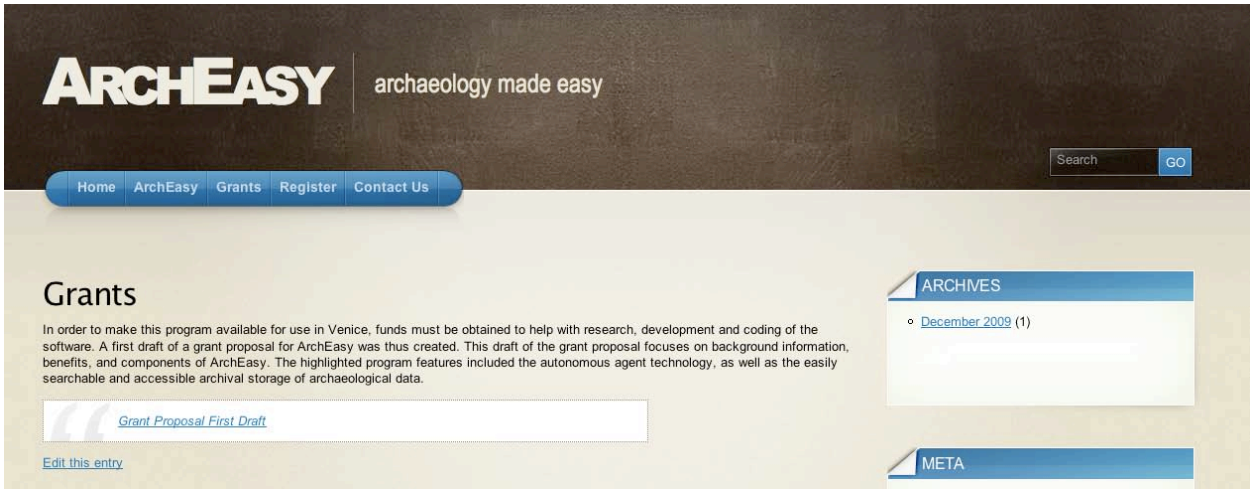


FIGURE 39: ARCHEASY GRANTS

APPENDIX C: WORCESTER POLYTECHNIC INSTITUTE IRB FORM

If your project has <u>any</u> federal sponsorship (e.g. federal funding), either prime or pass-through, the WPI IRB is <u>not authorized</u> to perform a review. Please contact Christina DeVries in Research Administration at (508) 831-6716 for direction to an appropriate IRB. <u>DO NOT</u> submit an application to the WPI IRB.

This application is for: *(Please check one)* ☒ Expedited Review ☐ Full Review

WPI IRB use only

Principal Investigator (PI) or Project Faculty Advisor: (*NOT a student or fellow; must be a WPI employee)*

Name:	Fabio Carrera	Tel No:	508-615-5333	E-Mail Address:	carrera@wpi.edu	☐
Department:	Interdisciplinary and Global Studies Division					

Co-Investigator(s): *(Co-PI(s)/non students)*

Name:	David Comas	Tel No:		E-Mail Address:	david.comas@upf.edu	☐
Name:		Tel No:		E-Mail Address:		☐

Student Investigator(s):

Name:	Andrew Kazanovicz	Tel No:	413-244-4953	E-Mail Address:	kaz@wpi.edu	☐
Name:	Benjamin Allen	Tel No:	508-612-4383	E-Mail Address:	allenbe@wpi.edu	☐

Check if: ☒ **Undergraduate project** *(MQP, IQP, Suff., other)*

☐ **Graduate project** *(M.S. Ph.D., other)*

Has an IRB ever suspended or terminated a study of any investigator listed above?

No ☒ Yes ☐ *(Attach a summary of the event and resolution.)*

Vulnerable Populations: The proposed research will involve the following (Check all that apply):

pregnant women ☐ human fetuses ☐ neonates ☐ minors/children ☐ prisoners ☐

students ☐ individuals with mental disabilities ☐ individuals with physical disabilities ☐

Collaborating Institutions: (*Please list all collaborating Institutions.)*

National Geographic and IBM - The Genographic Project

Locations of Research: (*If at WPI, please indicate where on campus. If off campus, please give details of locations.)*

Samples collected in Venice and surrounding areas, analysis completed at Universitat Pompeu Fabra, Barcelona, Spain

Project Title: Unearthing the Roots of Venice: From Relics to DNA

Funding: *(If the research is funded, please enclose one copy of the research proposal or most recent draft with your application.)*

Funding Agency: National Geographic and IBM – The Genographic Project WPI Fund:

Human Subjects Research: *(**All** study personnel having direct contact with subjects **must** take and pass a training course on human subjects research. There is a link to a web-based training course that can be accessed under the Training link on the IRB web site http://www.wpi.edu/Admin/Research/IRB/training.html. The IRB requires a copy of the completion certificate from the course or proof of an equivalent program.)*

Anticipated Dates of Research:

Instructions: Answer all questions. If you are asked to provide an explanation, please do so with adequate details. If attach itemized replies. Any incomplete application will be returned.

1.) Purpose of Study: *(Please provide a concise statement of the background, nature and reasons for the proposed nsert below using non-technical language that can be understood by non-scientist members of the IRB.)*

Start Date: 10/26/2009 Completion Date: 12/18/2009

2.) Study Protocol: *(Please attach sufficient information for effective review by non-scientist members of the IRB. Define all abbreviations and use simple words. Unless justification is provided this part of the application must not exceed 5 pages. Attaching sections of a grant application is not an acceptable substitute.)*

A.) For **biomedical, engineering and related research**, please provide an outline of the actual experiments to be performed. Where applicable, provide a detailed description of the experimental devices or procedures to be used, detailed information on the exact dosages of drugs or chemicals to be used, total quantity of blood samples to be used, and descriptions of special diets.

B.) For applications in the **social sciences, management and other non-biomedical disciplines** please provide a detailed description of your proposed study. Where applicable, include copies of any questionnaires or standardized tests you plan to incorporate into your study. If your study involves interviews please submit an outline indicating the types of questions you will include.

C.) If the study involves **investigational drugs or investigational medical devices**, and the PI is obtaining an Investigational New Drug (IND) number or Investigational Device Exemption (IDE) number from the FDA, please provide details.

D.) Please note if any **hazardous materials** are being used in this study.

E.) Please note if any **special diets** are being used in this study.

3.) Subject Information:

A.) Please provide the exact number of subjects you plan to enroll in this study and describe your subject population. *(eg. WPI students, WPI staff, UMASS Medical patient, other)*

Males: 200 Females: 0 Description: Venetian males over the age of 18

B.) Will subjects who do not understand English be enrolled?

No ☐ Yes ☒ *(Please insert below the language(s) that will be translated on the consent form.)*

C.) Are there any circumstances under which your study population may feel coerced into participating in this study?

No ☒ Yes ☐ *(Please insert below a description of how you will assure your subjects do not feel coerced.)*

D.) Are the subjects at risk of harm if their participation in the study becomes known?

No ☒ Yes ☐ *(Please insert below a description of possible effects on your subjects.)*

E.) How will subjects be recruited for participation? *(Check all that apply.)*

☒ Referral: *(By whom)* Fabio Carrera

☐ Other: *(Identify)* _____

☐ Database: *(Describe how database populated)* _____

F.) Have the subjects in the database agreed to be contacted for research projects? No ☐ Yes ☐ N/A ☐

☐ Direct subject advertising, including: *(Please provide a copy of the proposed ad. All direct subject advertising must be approved by the WPI IRB prior to use.)*

☐ Newspaper ☐ Bulletin board

☐ Radio ☐ Flyers

☐ Television ☐ Letters

☐ Internet ☐ E-mail

G.) Are the subjects being paid for participating? *(Consider all types of reimbursement, ex. stipend, parking, travel.)*

No ☒ Yes ☐ *(Check all that apply.)* ☐ Cash ☐ Check ☐ Gift certificate ☐ Other: _____

Amount of compensation _____

4.) Informed Consent:

A.) Who will discuss the study with and obtain consent of prospective subjects? *(Check all that apply.)*

☒ Principal Investigator ☐ Co-Investigator(s) ☐ Student Investigator(s)

B.) Are you aware that subjects must read and sign and Informed Consent Form prior to conducting any study-related procedures and agree that all subjects will be consented prior to initiating study related procedures? No ☐ Yes ☒

C.) Are you aware that you must consent subjects using only the IRB-approved Informed Consent Form? No ☐ Yes ☒

D.) Will subjects be consented in a private room, not in a public space? No ☐ Yes ☒

E.) Do you agree to spend as much time as needed to thoroughly explain and respond to any subject's questions about the study, and allow them as much time as needed to consider their decision prior to enrolling them as subjects? No ☐ Yes ☒

F.) Do you agree that the person obtaining consent will explain the risks of the study, the subject's right to decide not to participate, and the subject's right to withdraw from the study at any time? No ☐ Yes ☒

G.) Do you agree to either 1.) retain signed copies of all informed consent agreements in a secure location for at least three years or 2.) supply copies of all signed informed consent agreements in .pdf format for retention by the IRB in electronic form? No ☐ Yes ☒

(If you answer No to any of the questions above, please provide an explanation.)

5.) Potential Risks: *(A risk is a potential harm that a reasonable person would consider important in deciding whether to participate in research. Risks can be categorized as physical, psychological, sociological, economic and legal, and include pain, stress, invasion of privacy, embarrassment or exposure of sensitive or confidential data. All potential risks and discomforts must be minimized to the greatest extent possible by using e.g. appropriate monitoring, safety devices and withdrawal of a subject if there is evidence of a specific adverse event.)*

A.) What are the risks / discomforts associated with each intervention or procedure in the study?

Subject participation is just a cheek swab and a short questionnaire about family history, so very minimal risk/discomfort.

B.) What procedures will be in place to prevent / minimize potential risks or discomfort?

We will assure the test subject that a buccal swab is very quick and painless.

6.) Potential Benefits:

A.) What potential benefits other than payment may subjects receive from participating in the study?

Subject will get to know about their ancestry and help contribute to a better understanding of the origins of Venetians

---

B.) What potential benefits can society expect from the study?

Knowledge of origins of Venetians will add to the knowledge of the migratory paths and relations of mankind.

7.) Data Collection, Storage, and Confidentiality:

A.) How will data be collected?

Test subjects will be visited, informed consent will be signed and a cheek swab will be taken with a buccal brush.

B.) Will a subject's voice, face or identifiable body features *(eg. tattoo, scar)* be recorded by audio or videotaping?
No ☒ Yes ☐ *(Explain the recording procedures you plan to follow.)*

C.) Will personal identifying information be recorded? No ☒ Yes ☐ *(If yes, explain how the identifying information will be protected. How will personal identifying information be coded and how will the code key be kept confidential?)*

D.) Where will the data be stored and how will it be secured?

Data will be stored in the Genographic servers. Data is anonymous.

E.) What will happen to the data when the study is completed?

Data and results will be released to the public once the Genographic project is completed in 2010. All data is anonymous.

F.) Can data acquired in the study adversely affect a subject's relationship with other individuals? *(i.e. employee-supervisor, student-teacher, family relationships)*

No.

G.) Do you plan to use or disclose identifiable information outside of the investigation personnel?

No ☒ Yes ☐ *(Please explain.)*

H.) Do you plan to use or disclose identifiable information outside of WPI including non-WPI investigators?

No ☒ Yes ☐ *(Please explain.)*

8.) Deception: *(Investigators must not exclude information from a subject that a reasonable person would want to know in deciding whether to participate in a study.)*

Will the information about the research purpose and design be withheld from the subjects?

No ☒ Yes ☐ *(Please explain.)*

9.) Adverse effects: *(Serious or unexpected adverse reactions or injuries must be reported to the WPI IRB within 48 hours. Other adverse events should be reported within 10 working days.)*

What follow-up efforts will be made to detect any harm to subjects and how will the WPI IRB be kept informed?
None, but each participant will receive a number and access code so they can view their individual results online.

10.) Informed consent: (*Documented informed consent must be obtained from all participants in studies that involve human subjects. You must use the templates available on the WPI IRB web-site to prepare these forms.* ***Informed consent forms must be included with this application.*** *Under certain circumstances the WPI IRB may waive the requirement for informed consent.)*

Investigator's Assurance:

I certify the information provided in this application is complete and correct.

I understand that I have ultimate responsibility for the conduct of the study, the ethical performance of the project, the protection of the rights and welfare of human subjects, and strict adherence to any stipulations imposed by the WPI IRB.

I agree to comply with all WPI policies, as well all federal, state and local laws on the protection of human subjects in research, including:

- ensuring the satisfactory completion of human subjects training.
- performing the study in accordance with the WPI IRB approved protocol.
- implementing study changes only after WPI IRB approval.
- obtaining informed consent from subjects using only the WPI IRB approved consent form.
- promptly reporting significant adverse effects to the WPI IRB.

Signature of Principal Investigator _____ Date _____

Print Full Name and Title Dr. Fabio Carrera _____

Please return a signed hard copy of this application to the WPI IRB c/o Research Administration.

If you have any questions, please call (508) 831-6716.

Institutional Review Board Section 2.A

Unearthing the Roots of Venice: From Relics to DNA

The student IQP team will be performing the sample collection for the Genographic Project, which is a collaboration between National Geographic and IBM (see https://www3.nationalgeographic.com/genographic/). The Genographic contact is David Comas, who is one of the Principle Investigators of the Genographic Project. The Genographic Project aims to investigate the migratory history of mankind, by analyzing DNA samples from groups of people all over the world. The IQP team will be collecting cheek swab samples from Venetian males for analysis by the Genographic Project team. There are two types of genetic tests: mitochondrial DNA analysis (traces maternal lineage) and Y-chromosomal analysis (traces paternal lineage). Only male samples will be collected, since both maternal and paternal genetic tests can be performed on male DNA. Only maternal genetic tests can be performed on female DNA.

Test subjects will be contacted through Professor Fabio Carrera and the Venice Project Center. The subject will be asked if both sets of their grandparents are from the Venice region, confirming they are eligible for participation. The subject will be asked to swish their mouths with water. For sample collection, the following procedure will be used:

1. The swab brushes are individually packaged. Open the *Omni Swab* packaging at the handle end and carefully remove swab.
2. Holding the handle end of the swab, scrape the collection tip firmly against the inside of the cheek 5-6 times (about 10 seconds) being careful not to depress the plunger that ejects the tip.

3. After taking the sample, eject the tip by firmly pressing the plunger at the end of the handle into a labeled 2mL microcentrifuge tube (a small test tube).

The person administering the test will be wearing gloves. The test tube will be labeled with an identification code, and a code matching the sample will be given to the test subject. They can then go to the Genographic website and view their individual results. They can also withdraw their sample from the study/database if they choose to do so by using this code. The samples will be mailed to David Comas of the Universitat Pompeu Fabra in Barcelona, Spain.

APPENDIX D: DNA CONSENT FORM

The following are the procedural details of our Genographic Project collaboration.

INFORMED CONSENT FORM—ENGLISH

National Geographic Society in Collaboration with Worcester Polytechnic Institute

Title of Research Project: "*The Genographic Project: Molecular Genetic Analyses of Western/Central European populations*"

National Geographic Principal Investigators: Dr. Jaume Bertranpetit and David Comas, Unitat de Biologia Evolutiva, Universitat Pompeu Fabra. Doctor Aiguader 80 ,08003 Barcelona, Spain.

WPI Venice Project Center Project Title: *"Unearthing the Roots of Venice: From Relics to DNA"*

Overall Project Principal Investigator: Dr. Spencer Wells, Mission Programs, National Geographic Society, 1145 17th Street, N.W., Washington, D.C. 20036, 202-828-5465, SpWells@ngs.org

WPI Principle Investigator: Dr. Fabio Carrera, Worcester Polytechnic Institute, 100 Institute Rd, Worcester, MA 01609 USA. Email: carrera@wpi.edu Telephone: +39 041 523-3209

INFORMED CONSENT (Western/Central Europe)

Invitation to Participate: You are invited to participate in a project that studies the various human populations of Western/Central Europe through their inherited genetic properties. This project is sponsored by the National Geographic Society. The project team will talk to you about the study and how it might affect you. They will explain what you will have to do if you decide to participate. You can choose whether or not to participate. If you decide not to participate, no one will take any position of any kind against you.

Only unrelated males can participate, in order to maintain the validity of the study. In addition both maternal and paternal grandparents must have originated from the Veneto region.

Please read this consent form carefully. If you do not understand something, please ask the researcher to explain it. If you prefer, someone will read this to you. If you decide to participate, please sign the last page of this form. We will give you a copy to keep.

Purpose: This project is a major international effort to collect population genetic data from over 100,000 individuals around the world. Our sampling will target native human populations living preferably in their geographic region of origin and who ideally had minimal mixing with the surrounding populations. Our goal is to find out the relationships between genetic, linguistic, cultural and historical data. No data or information on your medical history will be collected from you, and thus the saliva sample collected with the buccal brush will never be used for any medically-related inquiry. To better understand the history of the native people from your region and around the world, the investigators will survey segments of your genetic fingerprint. Changes or variations in your genetic fingerprint may be considered as markers that can be used for comparison among populations. These changes are interesting for the study of origin and behavior of humans because many of them are found to group in descendants of a common ancestor within human populations. Because of the specific way of maternal and paternal inheritance, we can trace these genetic variations through human families from the present to the distant past with a

relatively high degree of accuracy. Therefore, by following the spread of the descendants of a specific ancestry, we can also genetically retrace the patterns of human movement through geographic areas.

The WPI Venice Project Center team is focusing on investigating the genetic heritage of the Venetian people in order to contribute to a better understanding of the origins of Venice. The data collected in the study will be used to help confirm/disprove different theories about the origins of the Venetian people.

Length of the Project: Your participation in the project will require one single visit taking 20 minutes of your time but the entire project will last 5 years. You will be one of many thousands of people in your region who give samples of blood or saliva.

Procedure: If you decide to participate, we will collect a cheek cell sample using a cotton brush. The saliva sample will be used to look for markers in your genetic fingerprint. These markers will be compared to those of ancient native populations. We will also ask questions about your family history.

Storage of Sample: Your saliva sample will be stored indefinitely at the Evolutionary Unit of the Universitat Pompeu Fabra (UPF), where it may be used for further study to better understand human origins and the histories of your people. The saliva sample cannot and will not be used for any medically related study. Furthermore, you may at any point in the future choose to have your sample removed from the project. You can do this by simply contacting our regional center (UPF, Doctor Aiguader 88, Tel. +34 93-3160843) and explaining your desire to have your sample removed. Your sample will then be immediately destroyed at the laboratory and all records and associated data from your sample will be eliminated. Instructions on how to do this are explained below.

Risks: You will spend about 20 minutes on the project. There is almost no risk associated with a cheek swab. Participation is voluntary and we cannot guarantee any results to you.

Benefits: The only benefits to you from participating in this project will be to learn more about your family origins and your relationships to other people around the world.

Costs: You do not have to pay anything to be in the project. You will not be paid for being in this project. If you have to travel to the collecting site, we may pay for your reasonable travel costs. We will not pay for your travel costs, *unless* you arrange for them to be paid for before you provide a blood sample.

Participation: The choice is yours. You may choose either to be in the project, or not to be in the project. Your participation is voluntary. There will be no penalty if you choose not to participate, or if you agree to be in the project but later change your mind. You may withdraw from the project at any time by calling or writing Dr. Jaume Bertranpetit or Dr. David Comas (+34 93-3160843) at the Unitat de Biologia Evolutiva, Universitat Pompeu Fabra, Doctor Aiguader 88, 08003 Barcelona, Spain, jaume.bertranpetit@upf.edu or david.comas@upf.edu. You may also withdraw from the project through the website at www.nationalgeographic.com/genographic **.** To withdraw from the project, you must provide your name and assigned code number. Also, the investigators may decide to discontinue your participation in this study without your permission.

Confidentiality: We will protect carefully the information that you tell us about yourself and your family. What we learn from your sample will be described only in a way that does not identify you. You will be given a code number for your sample and instructions how to use the code number to obtain the results of your sample. To protect your privacy, samples will be recorded with a secret code. Your name only will be recorded on the consent form. The secret code assigned to your sample will be kept in a locked file at the site where you participate in the project and carefully protected. No information related to your medical history will be included. Your sample will be stored at the regional site unless you ask to have your sample destroyed after the project. Your records will be monitored and may be audited without violating confidentiality.

SUBJECT STATEMENT OF CONSENT

I have read and understood the above description of this research study. All of my questions have been answered to my satisfaction. I know that my taking part in this research study is voluntary. I know that I may refuse to take part in or quit this research study at any time. If requested, a copy of this signed consent form will be given to me.

_____ _____

Signature of Subject Date

_____ _____

Signature of Witness Date

FOR STUDY USE ONLY

I attest that I have fully and appropriately informed the subject of the nature of the above research study and have offered to answer any question that he/she may have.

_____ _____

Signature of Principal Investigator/Designate Date

INFORMED CONSENT FORM—ITALIAN

Ricerca internazionale sulle origini degli Antichi Veneti attraverso lo studio del DNA

Una collaborazione tra Worcester Polytechnic Institute (WPI) e National Geographic Society

Nell'ambito del progetto internazionale: "*The Genographic Project: Molecular Genetic Analyses of Western/Central European populations"*
Direttore generale del progetto Genographic: Dr. Spencer Wells, Mission Programs, National Geographic Society, 1145 17th Street, N.W., Washington, D.C. 20036, 202-828-5465, SpWells@ngs.org

Ricercatore Principale di National Geographic: Dr. Jaume Bertranpetit and David Comas, Unitat de Biologia Evolutiva, Universitat Pompeu Fabra. Doctor Aiguader 80 ,08003 Barcelona, Spain.
Titolo della tesi WPI attualmente in corso: *"Unearthing the Roots of Venice: From Relics to DNA"*
Direttore del progetto WPI: Dr. Fabio Carrera, Worcester Polytechnic Institute, 100 Institute Rd, Worcester, MA 01609 USA. Email: carrera@wpi.edu Tel: +39 041 523-3209

INFORMAZIONI SULLA PRIVACY PER IL CONSENSO ALLA PARTECIPAZIONE

Invito a Partecipare: La invitiamo a partecipare ad un progetto che studia le diverse popolazioni umane dell'Europa Occidentale/Centrale tramite le loro caratterstiche genetiche. Questo studio è sponsorizzato dalla "National Geographic Society". Il gruppo di studenti che si occupa del progetto Le parlerà dello studio e di come potrebbe toccarLa. Può decidere liberamente se partecipare o meno al progetto e, qualora decisse di non partecipare, ciò non comporterà alcuna conseguenza nei Suoi confronti. Nel caso in cui decidesse di partecipare, Le verrà spiegato dettagliatamente che cosa deve fare.

Solo individui maschi non imparentati tra loro possono partecipare in modo di preservare la validità dello studio. Inoltre è indispensabile che il nonno e la nonna dei partecipanti siano di origini trivenete sia da parte di madre che di padre.

Per favore, legga attentamente questo documento. Se non dovesse comprenderne il contenuto o avesse delle domande al riguardo, chieda liberamente al ricercatore che sarà lieto di offirLe qualsiasi spiegazione. Se preferisce, qualcuno potrà leggerLe questo modulo. Se decisesse di partecipare, per favore firmi l'ultima pagina di questo documento. Le sarà fornita una copia da tenere per sè.

Descrizione dello Studio: Questo studio fa parte di un grande progetto internazionale che si prefigge di raccogliere dati genetici da oltre100.000 individui in tutto il mondo. I nostri campioni si concentrano su popolazioni indigene che vivono preferibilmente ancora nelle loro terre di origine e che preferibilmente si sono mescolate poco con altre popolazioni locali. Il nostro obiettivo è di scoprire le relazioni tra dati genetici, linguistici, culturali e storici. Non le verrà chiesto alcun dato medico nè informazioni sulla sua salute e quindi il campione genetico che le verrà prelevato non sarà assolutamente mai utilizzato per analisi igienico-sanitarie. I nostri ricercatori analizzeranno segmenti del suo DNA per comprendere meglio la storia delle popolazioni indigene della sua regione e del resto del mondo. Cambiamenti o variazioni nella sua impronta genetica verranno considerati come indicatori da utilizzare per confronti con altre popolazioni. Queste variazioni sono interessanti per lo studio delle origini e dei comportamenti delle stirpi umane dato che permettono di raggruppare popolazioni che discendono da antenati comuni. Grazie ai meccanismi di ereditarietà sia dal lato materno che paterno, possiamo ricostruire con discreta precisione queste variazioni genetiche attraverso stirpi umane dal presente fino alla piu remota antichità. Quindi, seguendo la distribuzione dei discendenti di una specifica

stirpe, possiamo ricostruire i percorsi migratori seguiti dalle varie popolazioni attraverso varie aree geografiche.
Lo studio specificamente condotto dal Venice Project Center del Worcester Polytechnic Institute nel contensto del piu ampio progetto *Genographics*, si concentra esclusivamente sul patrimonio genetico dei Veneti in generale e dei Veneziani in particolare, con l'intento di contribuire ad una migliore comprensione delle origini di Venezia. I dati raccolti da questo studio serviranno a confermare o smentire diverse teorie sulle origini geografiche dei Veneti, sulla loro diffusione nel continente europeo e soprattutto sul loro insediamento nelle lagune dove oggi sorge Venezia.

Durata dello Studio: La partecipazione allo studio comporterà un solo incontro con un addetto della durata di circa 20 minuti, mentre l'intero progetto avrà una durata totale di 5 anni. Lei sarà uno delle migliaia di persone della sua stessa regione a donare un campione di sangue o saliva.

Procedura: Qualora decidesse di partecipare, Le chiederemo di raccogliere un campione di saliva prelevandolo da sè dall'interno della guancia utilizzando uno strumento simile ad un piccolo spazzolino da denti con l'estremità tipo cottonfioc. Il campione verrà poi utilizzato per identificare alcuni marcatori specifici nella sua impronta genetica. Gli indicatori genetici così raccolti saranno comparati con quelli delle popolazioni native della regione. Le verrano inoltre richieste alcune brevi informazioni riguardanti la storia della Sua famiglia.

Stoccaggio del campione: Il Suo campione di saliva sara preservato indefinitamente dall'Unità Evoluzionaria dell'Università Pompeu Fabra (UPF) di Barcellona, dove potrebbe essere utilizzato per ulteriori studi per la comprensione delle origini storiche delle popolazioni mondiali. Ciononostante, sarà Sua facoltà di decidere di rimuovere il Suo campione dal progetto in qualsiasi momento futuro. Per far questo, basterà contattare il nostro centro regionale (UPF, Doctor Aiguader 88, Tel. +34 93-3160843) esprimendo il Suo desiderio di rimuovere il Suo campione dallo studio. Il Suo campione verrà allora immediatamente distrutto nel laboratorio e tutti i dati ad esso associati verranno anch'essi eliminati. Istruzioni dettagliate su questa procedura sono illustrate in calce.

Rischi, Inconvenienti e Disagi: La Sua partecipazione comporterà un impegno di tempo totale di circa 20 minuti. Al momento non si conoscono rischi per la salute nell'eseguire un prelievo di un campione di DNA. La partecipazione è volontaria e non ci è possibile garantirLe la buona riuscita del test genetico.

Benefici: L'unico reale beneficio derivante dalla partecipazione a questo studio sarà quello di avere maggiori informazioni riguardanti le origini della Sua famiglia e le relazioni tra questa e altre persone nel mondo.

Costi: Nessun costo è dovuto per la partecipazione al progetto. Non c'è alcuna ricompensa per la partecipazione.

Partecipazione: La Partecipazione a questo studio è volontaria. Puoi scegliere di non partecipare. Hai il diritto di ritirarti in qualsiasi momento e decidere se il tuo campione biologico, le informazioni genetiche o la genealogia già raccolti restino a far parte di questo studio o vengano distrutti.
Puoi decidere di ritirarti dal progetto in qualsiasi momento chiamando o scrivendo a Dr. Jaume Bertranpetit or Dr. David Comas (+34 93-3160843)alla Unitat de Biologia Evolutiva, Universitat Pompeu Fabra, Doctor Aiguader 88, 08003 di Barcelona, Spagna, jaume.bertranpetit@upf.edu oppure a david.comas@upf.edu. Può ritirarsi dal progetto anche attravero il sito web www.nationalgeographic.com/genographic. Qualora desiderasse ritirarsi dal progetto, dovrà comunicarci il suo nome e codice segreto identificativo assegnatoLe. Inoltre, i ricercatori del progetto potrebbero decidere di escludere il Suo campione dallo studio senza la Sua esplicita autorizzazione.

Riservatezza: La Sua identità, unica e personale, è considerata strettamente riservata e privata. La Sua identità, unica e personale, non sarà rivelata in alcuna pubblicazione generica o scientifica di dati. Campioni e documenti contenenti le Sue informazioni saranno tenuti in un luogo sicuro (vedere il paragraforelativo allo stoccagio del campione). Le uniche persone ad avere accesso ai codici e alle informazioni genealogiche saranno l'investigatore principale e le altre persone specificamente autorizzate dall'investigatore principale. La Sua identità unica e l'identità dei Suoi antenati recenti non sono collegate direttamente alle informazioni contenute nel database pubblico.
Le sarà assegnato un codice numerico e le istruzioni di utilizzo del codice stesso per poter accedere ai risultati relativi al Suo campione. a protezione della Sua privacy, il campione sarà quindi schedato con un codice segreto. Il Suo nome sarà registrato solo all'interno del presente Modulo di Consenso. Il codice segreto assegnato al Suo campione verrà attentamente custodito sotto chiave, all'interno di uno schedario appositamente protetto da intrusioni, presso la sede centrale del progetto. Nessuna informazione relativamente alla Sua situazione medica sarà inclusa. Il suo campione sarà custodito all'interno del sito locale relativo alla sua regione a meno che Lei non richieda esplecitamente la sua distruzione al termine dello studio. I suoi dati saranno sorvegliati ed occasionalmente visionati da addetti ai lavori senza che questo comporti la violazione della Sua privacy.

CONSENSO

Se è d'accordo a partecipare a questo studio, riceverà una copia firmata e datata di questo Modulo di Consenso per Suo uso personale.

Ho letto le informazioni su questo Modulo di Consenso. Ho ricevuto tutte le risposte alle mie domande in relazione allo studio e alla mia partecipazione. Sono consapevole che la mia partecipazione in questo studio di ricerca è frutto della mia volontà. Sono consapevole di avere il diritto di rifiutare di partecipare e anche di ritirarmi da questo studio in qualsiasi momento. Su richiesta, mi sarà fornita una copia di questo consenso firmato.

_____ _____
Firma del Soggetto Data

_____ _____

Firma del Testimone Data

AREA RISERVATA AI CONDUTTORI DELLO STUDIO

Attesto di avere completamente informato il soggetto circa la natura di questa ricerca e mi sono reso disponibile a rispondere a qualunque domanda che il soggetto avesse al riguardo.

_____ _____

Firma del Capo Ricercatore o del suo Delegato Data

Whatman Sterile Omni Swab
Instructions For Use

The Whatman Sterile ***Omni Swab*** is designed for collection of buccal cell samples for DNA testing. To obtain a buccal cell sample, the tip end of the ***Omni Swab*** is rubbed against the inside of the cheek. Unlike cotton-tipped swabs, the ***Omni Swab*** is made of absorbent material specifically designed for the collection of buccal cells. The single use swab has a unique removable collection tip to assist in processing the sample. While collecting a buccal sample is quick, easy and painless, it needs to be done correctly.

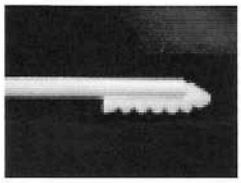
Suggested Directions for Use

The following directions are provided to illustrate an acceptable collection protocol. Each facility should establish a sample collection protocol that meets the objectives of collection and testing facilities.

Sample collection:

The person providing the buccal cell samples should not eat or drink immediately prior to giving the sample. If food or drink has been recently taken, it is suggested that the mouth be rinsed with water prior to sampling. The person taking the samples should wear biohazard-barrier gloves and avoid contact with the *Omni Swab* collection tip.

1. Open the *Omni Swab* packaging at the handle end and carefully remove swab. Do not touch the collection tip of the swab.
2. Holding the handle end of the *Omni Swab*, scrape the collection tip firmly against the inside of the cheek 5-6 times (about 10 seconds) being careful not to depress the plunger that ejects the tip.
3. After taking the sample, eject the tip by firmly pressing the plunger at the end of the handle into a labeled 2 mL microcentrifuge tube.
4. If desired, repeat sampling procedure with a second swab using the other cheek. Eject tip into a different labeled 2 mL microcentrifuge tube.



Storage:

After sample collection, tip can be kept at room temperature when processed immediately. If storage is necessary, freeze tips at -20°C.

DNA Extraction:

DNA can be extracted from the tip using standard laboratory procedures. DNA yields of 500-2,000 ng DNA are typical. Actual DNA yield will vary depending on the original DNA concentration, collection efficiency and extraction procedure.

Caution: Read instructions prior to use.
Single use only. Do not use if package has been previously opened or damaged. Federal law (USA) restricts this device to sale by or on the order of a physician.

To obtain Whatman® Sterile ***Omni Swabs***, Catalog #WB10-0004: Contact Whatman® toll free at 1-866-787-3362.

Whatman®
www.whatman.com

FIGURE 40. SWABBING INSTRUCTIONS INCLUDED IN DNA SAMPLE KIT

FIGURE 41. DNA SAMPLE COLLECTION KITS

APPENDIX F: PROJECT INFORMATION CARD

FIGURE 42. FRONT OF INFORMATION CARD

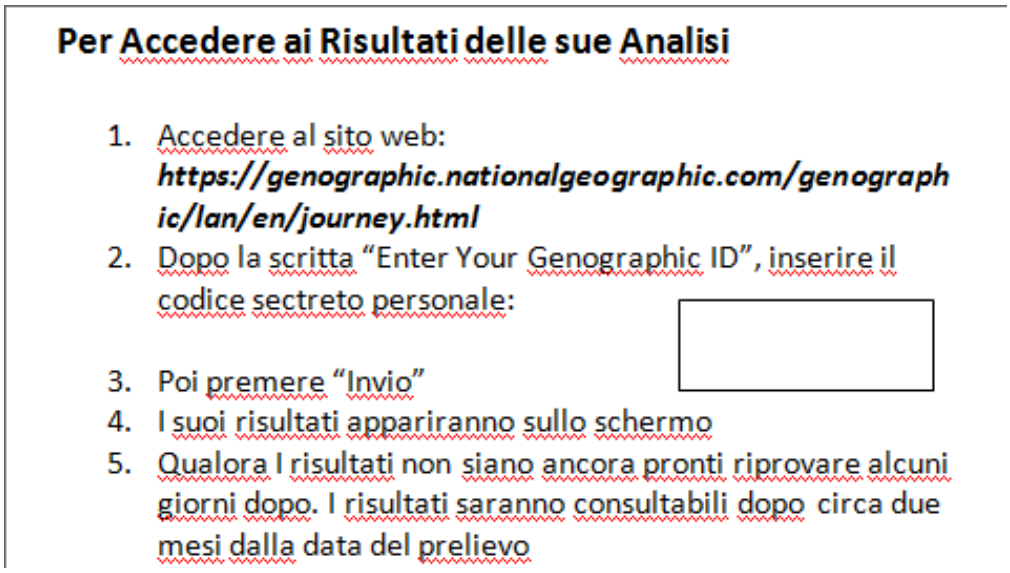


FIGURE 43. BACK OF INFORMATION CARD

APPENDIX G: ANCESTRY QUESTIONNAIRE

Sticker No.	Nonno Paterno	Nonna Paterna	Nonno Materno	Nonna Materna	Posta Elettronica (Ema
BA787*****	Venezia	Venezia	Venezia	Venezia	XXXXX
BA7J8*****	Burano	Venezia	Venezia	Venezia	XXXXX
BA4LD*****	Friuli	Venezia	Venezia	Venezia	XXXXX
BA54E*****	Venezia	Venezia	Venezia	Venezia	XXXXX
BA54N*****	Venezia	Venezia	Padova	Friuli	XXXXX
BARQM*****	Venezia	Friuli	Trau	Mantova	XXXXX
BARFJ*****	Venezia	Venezia	Reggine	Verona	XXXXX
BAUR7*****	Venezia	Chioggie	Friuli	Budio	XXXXX
BA549*****	Padova	Venezia	Venezia	Venezia	XXXXX
BASLN*****	Padova	Venezia	Murano	Venezia	XXXXX
BAH65*****	Venezia	Venezia	Venezia	Padova	XXXXX
BA689*****	Venezia	Udine	Padova	Padova	XXXXX
BA84A*****	Venezia	Friuli	Venezia	Trau	XXXXX
BA8AN*****	Venezia	Udine	Venezia	Friuli	XXXXX
BASPY*****	Venezia	Moventa	Venezia	Murano	XXXXX
BA8L7*****	Dolo-Venezia	Dolo-Venezia	Dolo-Venezia	Budio	XXXXX
BA9V8*****	Venezia	Venezia	Padova	Burano	XXXXX
BA659*****					
BAAD6*****					
BA5G5*****					

TABLE 4: SAMPLE ANCESTRY QUESTIONNAIRE

Identifying Genetic Traces of Historical Expansions: Phoenician Footprints in the Mediterranean

Pierre A. Zalloua,[1,2,13] Daniel E. Platt,[3,13] Mirvat El Sibai,[1] Jade Khalife,[1] Nadine Makhoul,[1] Marc Haber,[1] Yali Xue,[4] Hassan Izaabel,[5] Elena Bosch,[6] Susan M. Adams,[7] Eduardo Arroyo,[8] Ana María López-Parra,[8] Mercedes Aler,[9] Antònia Picornell,[10] Misericordia Ramon,[10] Mark A. Jobling,[7] David Comas,[6] Jaume Bertranpetit,[6] R. Spencer Wells,[11] Chris Tyler-Smith,[4,*] and The Genographic Consortium[12]

The Phoenicians were the dominant traders in the Mediterranean Sea two thousand to three thousand years ago and expanded from their homeland in the Levant to establish colonies and trading posts throughout the Mediterranean, but then they disappeared from history. We wished to identify their male genetic traces in modern populations. Therefore, we chose Phoenician-influenced sites on the basis of well-documented historical records and collected new Y-chromosomal data from 1330 men from six such sites, as well as comparative data from the literature. We then developed an analytical strategy to distinguish between lineages specifically associated with the Phoenicians and those spread by geographically similar but historically distinct events, such as the Neolithic, Greek, and Jewish expansions. This involved comparing historically documented Phoenician sites with neighboring non-Phoenician sites for the identification of weak but systematic signatures shared by the Phoenician sites that could not readily be explained by chance or by other expansions. From these comparisons, we found that haplogroup J2, in general, and six Y-STR haplotypes, in particular, exhibited a Phoenician signature that contributed > 6% to the modern Phoenician-influenced populations examined. Our methodology can be applied to any historically documented expansion in which contact and noncontact sites can be identified.

The Phoenicians were a distinctive and independent civilization that dominated the Mediterranean Sea during the first millennium BCE, emerging from a coastal section of the Eastern Mediterranean, including the four main Bronze Age maritime cities of Tyre, Sidon, Byblos, and Arwad and located in the modern countries of Lebanon and southern Syria. From here, their maritime expertise allowed them to establish a trading empire throughout the Mediterranean and beyond.[1–6] Their strategy included the establishment of settled colonies, foremost among which was Carthage in modern Tunisia, and many trading posts, where they stayed for shorter periods[4] (Figure 1A). Their activities were recorded by contemporary writers, including the Egyptians, the Greeks, Biblical sources, Strabo, Pliny the Elder, and Avienus, and the remains of their cities and trading goods have been documented extensively by archaeologists.[6] Thus, we have a good understanding of their origins and spread from historical sources.

We set out to complement this historical information by searching for Phoenician genetic traces within modern populations. We chose the nonrecombining region of the Y chromosome for this purpose, because its male specificity means that it would have been carried by the predominantly male Phoenician traders, and its high level of geographical differentiation provides the greatest chance of recognizing colonization events.[7] Human genetic history, however, can be viewed as a palimpsest, in which multiple events from different times but with similar geographical patterns are superimposed. Expansions from the Eastern Mediterranean could include the initial peopling by modern humans during the Paleolithic era, the subsequent Neolithic-era transition originating in the Fertile Crescent ~8000 BCE, and later events, such as the Greek expansion or the Jewish Diaspora. All of these, and possibly additional events unrecorded in history, could result in broadly similar genetic patterns with an origin in or near the Levant and decreasing gradients toward the west. Several previous studies have identified Y-chromosomal types showing gradients originating in the Near East[8–11] and have sometimes linked them to the Phoenicians,[12] but further work is needed to distinguish between the general pattern and the specific Phoenician contribution.

Therefore, we have developed a strategy for identifying a geographical genetic pattern associated with a specific historical expansion, rather than an overall geographical gradient. The key to this was the use of historically documented locations of greater or lesser contact—in our case, Phoenician locations—matched approximately for

[1]The Lebanese American University, Chouran, Beirut 1102 2801, Lebanon; [2]Harvard School of Public Health, Boston, MA 02215, USA; [3]Bioinformatics and Pattern Discovery, IBM Thomas J. Watson Research Centre, Yorktown Heights, NY 10598, USA; [4]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK; [5]Laboratoire de Biologie Cellulaire & Génétique Moléculaire, Université Ibn Zohr, 8106 Agadir, Maroc; [6]Institute of Evolutionary Biology, Consejo Superior de Investigaciones Científicas, Parc de Recerca Biomèdica de Barcelona, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, 08003 Barcelona, Catalonia, Spain; [7]Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK; [8]Universidad Complutense de Madrid, Facultad de Medicina, Ciudad Universitaria, 28040 Madrid, Spain; [9]Unidad Docente de Medicina Legal, Sección de Biología Forense, Facultad de Medicina, Universidad de Valencia, 46010 Valencia, Spain; [10]Laboratori de Genètica, Institut Universitari d'Investigació en Ciències de la Salut i Departament Biologia, Universitat de les Illes Balears, 07122 Palma de Mallorca, Spain; [11]The Genographic Project, National Geographic Society, Washington, DC 20036, USA; [12]Genographic Consortium members are listed fully in the Acknowledgments section
[13]These two authors contributed equally to this work
*Correspondence: cts@sanger.ac.uk
DOI 10.1016/j.ajhg.2008.10.012.

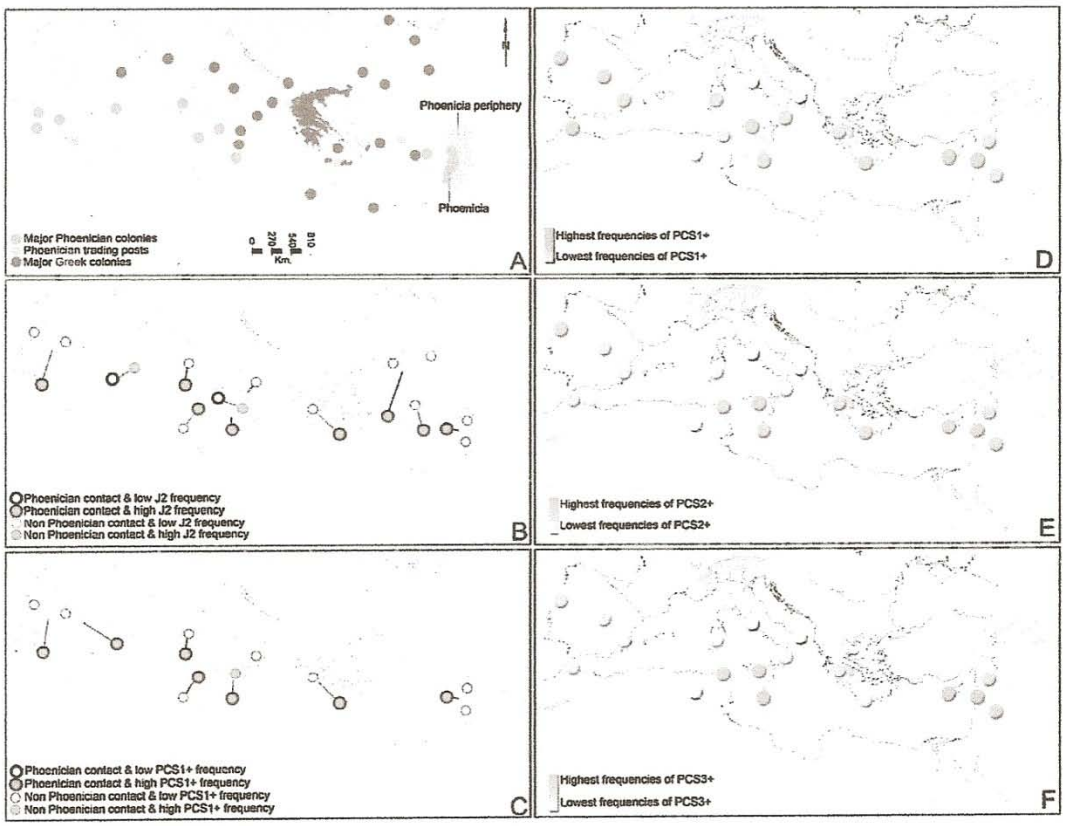


Figure 1. Geographical context of the Phoenician and Greek expansions

(A) Maritime expansions of the Phoenicians (11th century BCE) and Greeks. Red: Phoenicia, Phoenician colonies; pink: Phoenician trading posts; blue: Greece and Greek colonies.

(B) J2 haplogroup frequency comparisons between Phoenician contact regions (thick borders) and nearby non-contact regions (thin borders). Lines indicate paired haplogroup comparisons between two sites. An ellipse indicates a site with multiple population samples. Colored circles indicate the higher haplogroup J2 frequency site in each pair.

(C) Phoenician Colonization Signal 1 (PCS1+) haplotype frequency comparisons between Phoenician contact regions (thick borders) and nearby non-contact regions (thin borders). Lines indicate paired haplotype comparisons between two sites. An ellipse indicates a site with multiple population samples. Colored circles represent the higher PCS1+ frequency site in each pair.

(D–F) Geographical distribution of the PCS1+ (D), PCS2+ (E), and PCS3+ (F) haplotypes in the Mediterranean region. The PCS+ central haplotypes are shown in Table 2. Higher color intensities indicate higher haplotype frequencies; absolute frequencies are given in Table 3. Note the highly enriched coastal and island distribution of these haplotypes and the prominence of all in the Levant.

distance from the source. Such paired locations would be expected to share general genetic patterns, reflecting the sum of multiple events, but to differ specifically in their Phoenician genetic influence if genetic transfer had taken place. Other historical expansions would have involved different locations of greater and lesser contact and so would not have produced a distinct geographically detailed signal in the same populations at this fine level of resolution. To assess the significance of any pattern that we might detect, we developed a two-fold analytic approach: first, a statistical component, the investigation of whether such a pattern might have originated by chance; and second, an empirical component, the application of the same analytical strategy to additional data sets not expected to differ in their Phoenician influence, representing instead the general Neolithic spread or the independent Greek expansion. Haplotypes that would not be expected to exhibit the specific short-ranged variational features by chance and that did not correspond to other known expansions could be considered as potentially Phoenician. With the very active intervening history, we cannot reasonably expect to identify a statistically significant signature linking the Phoenician homeland to every identified colonization region. However, colonization is expected to

have produced a noisy but *systematic* trace of signatures. This study presents a method that identifies significant geographical preponderance of such signatures in order to decipher the genetic palimpsest.

In order to apply this strategy, we therefore needed to (1) choose suitable population sample sites for investigating Phoenician and other expansions, (2) generate or identify from available sources Y-chromosomal data sets from the chosen sites, (3) develop our test methodology, and, finally, (4) consider the broader significance of any signals that emerged from the chosen population sites.

When choosing populations, we considered that trade-driven colonization would have mediated the genetic legacy of the Phoenician expansion. Minor colonization sites were established for the servicing of ships en route, as well as for connecting with and guarding interests in foreign trade centers. This servicing was necessary for the expansion of trade throughout the Mediterranean basin with the maritime technology of the first millennium BCE and established the regional variations that we seek to detect. Carthage emerged as the dominant Central Mediterranean colony, connecting western-metals trade to the rest of the wealthy Mediterranean trading sites. Opportunity for establishing Phoenician colonization was greatest and most lasting in minimally occupied regions. Documented major colonies and trading posts are summarized in Figure 1A. We constructed pairs of testing sites generally orthogonal to the anticipated background of the Neolithic gradient originating in the Levant, resulting in localized groups of tests. The Phoenician-influenced regions selected were, thus, the coastal Lebanese Phoenician Heartland and the broader area of the rest of the Levant (the "Phoenician Periphery"); then Cyprus and South Turkey; then Crete; then Malta and East Sicily; then South Sardinia, Ibiza, and Southern Spain; and, finally, Coastal Tunisia and cities like Tingris in Morocco (Figures 1B and 1C). For each, we identified nearby sites of lesser or no Phoenician contact. Examples of the comparisons used thus include heartland versus periphery, colony versus trading center, and trading center versus noncontact sites.

In addition, we sought to discriminate Phoenician candidate lineages from those spread by other colonization expansions affecting many of the same islands and regions. We constructed a Neolithic-expansion test set by choosing paired sites from the region, both of which lacked known Phoenician contact, and comparing the site closer to the Levant with that farther away (Table 1 and Table S3, available online). The colonization by Greeks and later groups occurred largely into regions still unoccupied by the Phoenicians, yielding colonial segregation; Greek sites are also shown in Figure 1A. We wished to design similar tests to evaluate a potential signature of the Jewish Diaspora. This, however, proved problematic. At the time of the Roman destruction of Herod's Temple in 70 CE, there were already more Jews living outside than within Israel.[13] The dispersals through time and space were complex, with communities being established and dispersed, sometimes on multiple occasions. It was, thus, difficult to identify any locality where significant Jewish settlement did not occur for at least some period.[14] Therefore, our approach was not suitable for identifying lineages associated with the Jewish Diaspora, which has already been well studied with the use of other approaches.[15]

Data from Lebanon were available,[16] and we collected 1330 additional DNA samples from Syrian, Palestinian, Tunisian, Moroccan, Cypriote, and Maltese males with at least three generations of indigenous ancestry. Each provided information on their geographical origin and gave informed consent for this study. Samples were typed with 11 STRs and with 58 Y-SNPs as described elsewhere[16] (Table S1). We augmented our collection with suitable published data on 5,899 males from 56 sites (Table S2). Desirable sites that we were unable to include in our analysis included Libya and southern France, both of which could have provided more Greek coastal-settlement sites. The Y-chromosomal data were of two types: haplogroup data based on Y-SNPs and haplotype data based on Y-STRs. Although both types are carried on the same chromosome and are correlated,[17] they were analyzed separately, because they have different mutational properties and because some data sets contain only one of the two data types. A reduced set of haplogroups that captured most of the SNP information was used as previously.[16] It was also necessary to develop a similar procedure for the STR information by enumerating the regions and sizes of samples captured by various combinations of STR subsets, through a process informed by association-discovery methods.[18] We identified subsets containing seven STRs that maximized regional coverage and sample number, yielding the STRs DYS19, DYS389I, DYS389b (consisting of DYS389II–DYS389I), DYS390, DYS391, DYS392, and DYS393. We lost STR coverage of some regions, reducing the number of tests that were applied to the STR set. The geographical coverage of the STR samples and the SNP samples was not identical, and the regional tests that could be constructed from historical references were not identical for both genetic marker types. For example, Moroccan samples were included and tested in the STR set but not in the SNP-typed set.

The tests were constructed and validated in several ways. First, a noncontact test-pair matrix was constructed for detecting general east-to-west background variation reflecting Neolithic migrations, and the data were evaluated for significant results reflecting general non-Phoenician background variation against which the Phoenician pattern must be identified. Second, a colonization test-pair matrix for identification of gross features of the subsequent and more widespread Greek colonization event was applied. The Greek test sought to identify features typical of the Greek expansion but focused on those characteristics distinct from the Phoenician expansion. Third, the Phoenician colonization of Tunisia presented a unique test between the colonized coastal regions and interior Berber and Arab populations, because it has a different Neolithic history[19] and no intervening Greek-colonization events.

Table 1. Y-SNP Haplogroup Colonization-Site Gradient Tests with Aggregate Scores for Phoenician Colonies, Neolithic Background[a], and Greek Colonies

Tests	E3b	G	I	J*(xJ2)	J2	K2	L	R1a	R1b
Phoenician Test Sites									
Heartland versus Periphery	0.574	0.986	0.012	1.000	0.011	0.137	0.0002	0.894	0.003
	−1	−1	+1	−1	+1	+1	+1	−1	+1
Homeland versus Levant	0.968	0.833	0.260	1.000	0.000	0.188	0.0002	0.033	0.001
	−1	−1	+1	−1	+1	+1	+1	+1	+1
Cyprus versus S. Turkey	0.249	1.0	0.337	0.963	0.150	0.586	–	0.957	0.973
	+1	−1	+1	−1	+1	+1	–	−1	−1
S. Turkey versus N. Turkey	0.278	0.860	0.794	0.233	0.449	0.541	1.000	0.194	0.417
	+1	−1	−1	+1	+1	+1	−1	+1	+1
Lowland Crete versus Lasithi Plateau	0.211	0.685	0.436	0.858	0.000	0.905	–	0.988	1.000
	+1	−1	+1	−1	+1	−1	–	−1	−1
Crete versus Greece	0.994	0.783	0.989	0.624	0.0003	0.338	–	0.897	0.145
	−1	−1	−1	+1	+1	+1	–	−1	+1
Malta versus Sicily	1.000	0.561	0.434	1.000	0.016	1.000	–	0.083	0.653
	−1	0	+1	−1	+1	−1	–	+1	−1
W. Sicily versus E. Sicily	0.962	0.035	0.0711	0.0893	0.973	0.814	–	0.666	0.131
	−1	+1	+1	−1	−1	−1	–	0	+1
Sicily versus S. Italy	0.816	0.936	0.376	0.573	0.208	0.570	–	0.788	0.692
	−1	−1	+1	+1	+1	+1	–	−1	+1
S. Sardinia versus N. Sardinia	0.736	0.935	0.123	0.141	0.206	–	–	0.677	0.561
	−1	−1	+1	+1	+1	–	–	+1	−1
Ibiza versus Mallorca & Minorca	0.846	0.046	0.956	1.000	0.842	0.000	–	1.000	0.941
	−1	+1	−1	−1	−1	+1	–	−1	−1
S. Spain versus Valencia	0.767	0.317	0.896	0.738	0.142	1.000	–	0.738	0.539
	−1	+1	−1	−1	−1	−1	–	−1	+1
Contact Spain versus Iberia	0.879	0.988	0.259	0.807	0.176	0.385	–	0.141	0.197
	−1	−1	+1	0	+1	+1	–	+1	+1
Coastal Tunisa versus Inland Tunisia	0.890	–	–	–	0.0013	0.863	–	1.000	0.952
	−1	–	–	–	+1	−1	–	−1	−1
$\alpha = 0.05$	1.000	0.135	0.486	1.000	3.3×10^{-5}	0.486	0.0073	0.512	0.153
$\alpha = 0.30$	0.839	0.936	0.579	0.798	2.5×10^{-4}	0.420	0.216	0.839	0.644
Δf	0.993	0.987	0.133	0.927	0.0287	0.290	0.5	0.910	0.395
Control Test Sites									
Turkey #5 versus Turkey #3	0.555	0.691	0.421	0.794	0.275	0.421	1.000	0.148	0.557
	+1	−1	+1	+1	+1	+1	−1	+1	+1
Turkey #8 versus Turkey #1	0.941	0.0252	0.866	0.987	0.982	0.366	0.601	0.750	0.315
	−1	+1	−1	−1	−1	+1	+1	−1	+1
Greece versus Albania	0.578	0.655	0.988	0.944	0.735	0.605	0.605	0.499	0.157
	0	+1	−1	−1	−1	+1	+1	+1	+1
Serbia versus Croatia	0.013	0.724	1.000	–	0.022	–	–	0.484	0.181
	+1	+1	−1	–	+1	–	–	+1	+1
Italy WCL versus Italy NWA	0.058	0.760	0.553	0.580	0.036	–	–	0.963	0.990
	+1	−1	+1	+1	+1	–	–	−1	−1
Italy WCP versus Italy CMA	0.007	0.285	0.316	0.602	0.987	–	–	0.570	0.725
	+1	+1	+1	+1	−1	–	–	+1	−1
Italy SLA versus Italy NEL	0.999	0.671	0.458	–	0.121	–	–	0.471	0.617
	−1	−1	+1	–	+1	–	–	+1	−1
Italy TLB versus Italy EBL	0.257	0.244	0.999	–	0.034	–	–	0.840	0.960
	+1	+1	−1	–	+1	–	–	−1	−1
S. Portugal versus N. Portugal	0.998	0.377	0.601	0.031	0.223	–	–	0.246	0.583
	−1	+1	0	+1	+1	–	–	+1	−1
S. Greece versus N. Greece	0.229	0.511	0.071	0.974	0.418	0.694	1.000	1.000	0.3000
	+1	+1	+1	−1	+1	0	−1	−1	+1
S. Egypt versus N. Egypt	0.984	0.164	0.402	0.153	0.931	0.176	–	1.000	0.409
	−1	+1	+1	+1	−1	+1	–	−1	+1
$\alpha = 0.05$	0.102	0.432	0.432	0.337	0.015	1.000	1.000	1.000	1.000
$\alpha = 0.30$	0.210	0.828	0.980	0.748	0.078	0.760	1.000	0.887	0.887
Δf	0.377	0.113	0.377	0.363	0.274	0.688	0.688	0.500	0.500

Table 1. *Continued*

Tests	E3b	G	I	J*(xJ2)	J2	K2	L	R1a	R1b
Greek Test Sites									
Greece & Crete versus Turkey #7	0.1951	0.9264	0.0183	0.9865	0.4270	0.5180	1.0000	0.0513	0.5364
	+1	−1	+1	−1	+1	+1	−1	+1	+1
Greece & Crete versus Turkey #4	0.7091	0.6836	0.0000	1.0000	0.3777	0.1891	1.0000	0.4615	0.3583
	−1	−1	+1	−1	+1	+1	−1	+1	+1
E. Sicily versus Sardinia	0.0002	0.9994	1.0000	0.0272	0.0000	0.0032	–	0.1462	0.5451
	+1	−1	−1	+1	+1	+1	–	+1	+1
E. Sicily versus W. Sicily	0.0735	0.9904	0.9779	0.2649	0.0552	0.4294	–	0.6833	0.9271
	+1	−1	−1	+1	+1	+1	–	−1	−1
Greece & Crete versus Cyprus	0.9410	0.0096	0.1852	0.9881	0.864	0.717	–	0.0184	0.0184
	−1	+1	+1	−1	+1	−1	–	+1	+1
Greece & Crete versus S. Italy	0.9967	0.9712	0.0815	0.9572	0.0110	0.493	–	0.0144	0.0144
	−1	−1	+1	−1	+1	+1	–	+1	+1
Greece & Crete versus Spain	0.2539	0.0937	0.00142	0.9252	0.000	0.0000	–	0.0000	0.0000
	+1	+1	+1	−1	+1	+1	–	+1	+1
Turkey #8 & #9 versus Lebanon	0.9967	0.0008	0.5830	1.0000	0.7884	0.906	0.97	0.1899	0.1899
	−1	+1	−1	−1	−1	−1	−1	+1	+1
Turkey #8 & #9 versus Palestinian	0.9998	0.0218	0.3905	1.0000	0.0310	0.711	0.2133	0.0195	0.0195
	−1	+1	+1	−1	+1	−1	+1	+1	+1
Greece versus Lebanon	0.1653	0.1351	0.0000	1.0000	0.979	0.788	1.0	0.0000	0.0000
	+1	+1	+1	−1	−1	−1	−1	+1	+1
Greece versus Palestinians	0.5177	0.3013	0.0000	1.0000	0.4529	0.602	1.0	0.0000	0.0000
	+1	+1	+1	−1	+1	+1	−1	+1	+1
Crete versus Lebanon	0.9908	0.6702	0.0044	1.0000	0.0003	0.312	1.0	1.0	0.0000
	−1	−1	+1	−1	+1	+1	−1	−1	+1
Crete versus Palestinians	0.9993	0.8484	0.0043	1.0000	0.0000	0.142	1.0	0.0000	0.0000
	−1	−1	+1	−1	+1	+1	−1	+1	+1
S. Italy versus Coastal Tunisia	0.9358	0.0111	0.1771	0.1771	0.4996	0.7255	–	0.4253	0.0026
	−1	+1	+1	+1	+1	−1	–	+1	+1
Sicily versus Coastal Tunisia	0.9889	0.0430	0.0597	0.1834	0.1989	0.6332	–	0.4672	0.0018
	−1	+1	+1	+1	+1	+1	–	+1	+1
$\alpha = 0.05$	0.5367	0.0006	0.0000	0.5367	0.0000	0.1710	1.0000	0.0000	0.0000
$\alpha = 0.30$	0.7031	0.1311	0.0007	0.7031	0.0500	0.7031	0.9423	0.0037	0.0006
Δf	0.8491	0.5000	0.0176	0.9824	0.0037	0.1509	0.9961	0.0037	0.0005

[a] Designated control.

Additionally, the Moroccan military colonies are expected to be weaker than the major Phoenician Tunisian-trade-based colony but also to lack Greek influence.

Test-site pairs for haplogroups relevant to the Phoenician expansion are indicated in Figure 1B, and those for STR-defined haplotypes are indicated in Figure 1C. Preponderance p values representing test-pair aggregates were computed as described below, and two techniques were employed to establish these measures.

The first test was direct frequency comparison by means of the binomial sign test. Test sites were scored as positive if the contact-site frequency was larger than the noncontact-site frequency. The number of positive results, N_+, out of a total of N tests expected by chance should be randomly distributed following a binomial distribution with p = 0.05, so the probability that N_+ or more would have been observed by chance according to the "nonparametric" binomial sign test is

$$p_{\geq N_+} = \sum_{n=N_+}^{N} \binom{N}{n} p^n (1-p)^{N-n} = \sum_{n=N_-}^{N} \binom{N}{n} 2^{-N}.$$

Second, we applied Fisher's exact test to determine the chances of drawing m_+ or more out of t by chance given that they were taken randomly from M_+ total stronger contact samples and M_- total weaker contacts across the two test regions, with probability

$$p_{\geq}(m_+) = \sum_{m=m_+}^{M_+} \binom{M_+}{m}\binom{M_-}{t-m} \Big/ \binom{M_+ + M_-}{t}.$$

By the probability-integral-transform theorem, the distribution of p values may itself be considered to be a uniformly distributed random variable over the interval [0,1]. At a confidence level of α, the site was considered a positive candidate if $p_{\geq}(m_+) \leq \alpha$. This would be expected to be satisfied an α fraction of the time. For an individual site, a significant ($\alpha = 0.05$) or highly significant ($\alpha = 0.01$) level is usually required. However, testing for randomness even with much larger values of α is possible for putatively independent sites with the use of the binomial test, in the same way that the fairness of dice or the fairness of a coin may be tested.[20] Then the probability of seeing N_+ or more sites by chance at significance level

of α can also be tested according to the binomial test, such that $p_{\geq N_+}(\alpha) = \sum_{n=N_+}^{N} \binom{N}{n} \alpha^n (1-\alpha)^{N-n}$.[20] Even for a relatively weak α level of significance, the probability of seeing multiple sites at that level can yield a highly significant preponderant probability. Fisher's exact test tends to be more demanding for small samples, and if the sample is too small, it will never yield significant results. Yet, the number of times that relatively small samples will satisfy a weak significance of, say, $\alpha = 0.30$ still provides opportunity for probing the significance of sites with such small samples and for counting their contribution in determining the overall probability of seeing a Phoenician signal by chance.

Because there is a significant chance that a haplotype existing 3000 years ago has accumulated a one-step difference in an STR (we expect 0.6 mutations per seven-STR haplotype when a rate of 6.9×10^{-4} per locus per 25 yr is used[21]), these one-step neighbors have been included in each set, producing what we have labeled STR+s. STR+s can contain both haplotypes deriving from mutations, which should have been included, and independent haplotypes unconnected with the migrations that we are trying to detect. Those other sources are expected to be uncorrelated and incoherent relative to the signals we seek. STR sets can be found within multiple haplogroups, so contributions from multiple haplogroups might contribute to each of the STR+ samples as well, providing further stochastic background noise. Among STR+s, test sites were excluded when gradient differences were computed if there were two or fewer total STR+ samples in both sites. If there were fewer than three total STR+ samples, the Fisher's exact probabilities were discarded, because many comparison configurations can never show significant probabilities with such small samples.

The number and relative frequency of the major haplogroups observed in the sample regions employed in this study are shown in Table S3. Table 1 represents the outcome of Phoenician-colonization tests, the Neolithic control tests, and the Greek-colonization tests. Each of the results shows the Fisher's exact test p value as a number between 0 and 1, together with frequency-difference test as +1 or −1. Aggregate scores computed on the Fisher's exact test results for thresholds $\alpha = 0.30$ and $\alpha = 0.05$, as well as for counts of Δf signs of frequency differences, are reported at the end of each section. The $\alpha = 0.05$ results measure whether the number of strongly different gradients is significantly different than that expected by chance, whereas the $\alpha = 0.30$ results reports the same for modestly preferential sites, identifying a persistent pattern of weaker signals. Although any individual signal at this lower significance level might not be significant, the signal across all sites could be. The frequency-differences test Δf seeks to report signal in cases in which the number of samples may be low but may still contribute to a preponderance of evidence.

The Neolithic control section shows nonsignificant results across all haplogroups, except for a significant J2 result in one test. The Phoenician-colony test results highlight only one haplogroup, J2, which consistently scores significantly in all three tests across the range of colonization sites (Table 1, Figure 1B). However, this haplogroup also scores significantly in Greek tests (as do some additional haplogroups; Table 1), suggesting that the same haplogroup could have been spread by several expansions, which is unsurprising considering its frequency in the Eastern Mediterranean but implies that higher phylogenetic resolution is required for identification of Phoenician-specific signals.

Table 2. Core Haplotypes Defining Y-STR Haplotype Groups[a] Associated with the Phoenician or Greek Expansions

STR+	DYS19,DYS389I,DYS389b,DYS390, DYS391,DYS392,DYS393
PCS1+	14,13,16,24,10,11,12
PCS2+	14,14,17,23,10,11,12
PCS3+	13,12,18,23,10,11,13
PCS4+	14,13,17,23,10,11,12
PCS5+	14,14,16,23,10,11,12
PCS6+	14,13,16,22,10,11,12
GCS1+	13,13,17,24,10,11,13

[a] Designated STR+s.

Table 2 shows the core STR haplotypes of the STR+ groups that we focus on, and Table S4 reports the population frequencies for these STR+s. These STR+ groups were labeled "Phoenician Colonization Signal" or PCS1+ through PCS6+. Among the total of 1268 STR+s identified, 1237 showed coverage at nine or fewer sites. From the remaining 31, several candidates—PCS1+, PCS2+, PCS4+, PCS5+, and PCS6+—were identified from their high p values (Table 3). PCS3+ scores strongly as a Phoenician-colonization candidate and is strongly associated with the SNP haplogroup E3b, but it does not show the wide geographic coverage that the other PCS+s demonstrate. It represents the strongest of the lower-coverage STR+s. Both PCS1+ and PCS2+ score well, although not as strongly as PCS3+. However, they show much broader penetration throughout the Mediterranean, and both score relatively weakly as Greek-colonization candidates. Of these, PCS1+ shows a nearly significant Greek score for $\alpha = 0.05$ because of low representation in Tunisia, but it shows significant representation in Morocco, and the Greek score is simply an artifact. Both PCS1+ and PCS2+ contain multiple haplogroups, primarily J2 but including J*(xJ2) and E3b, with PCS1+ containing the greatest diversity. The use of one-step STR+s provides the opportunity to pick up mutated descendants of those who participated in colonization, as intended, but also of those who acquired the same signature through alternative paths. As expected, those other paths have tended to degrade a systematic colonization signal, shown by the relatively weak $\alpha = 0.05$ score relative to $\alpha = 0.30$ in comparison with PCS3+. A "Greek Colonization Signal" STR+ group, GCS1+, was also identified, which scored weakly as a Phoenician

Table 3. STR+ Colonization Site Gradient Tests, with Aggregate Scores for Phoenician Colonies and Greek Colonies

STR+ Tests	PCS1+	PCS2+	PCS3+	PCS4+	PCS5+	PCS6+	PCS1+ through PCS3+	GCS1+
Phoenician Test Sites								
Phoenician Heartland versus Phoenician Periphery	0.425	0.609	0.922	0.819	0.467	0.098	0.606	0.725
	+1	−1	−1	−1	−1	+1	−1	−1
Phoenician Heartland versus Palestinians	0.000	0.078	0.370	0.156	0.004	0.014	0.000	0.983
	+1	+1	+1	+1	+1	+1	+1	−1
Phoenician Periphery versus Palestinians	0.000	0.079	0.080	0.042	0.012	0.272	0.000	0.943
	+1	+1	+1	+1	+1	+1	+1	−1
Syria versus Palestinians	0.241	0.034	–	0.858	0.247	0.969	0.022	0.961
	+1	+1	+1	−1	+1	−1	+1	−1
Cyprus versus N and S Turkey	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–
S Turkey versus N Turkey	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–
Crete versus Greece	0.093	0.194	–	0.030	0.257	0.016	0.040	0.876
	+1	+1	–	+1	+1	+1	+1	−1
Malta versus W Sicily and S Italy	0.498	0.029	–	0.005	–	–	0.063	1.000
	+1	+1	–	+1	–	–	+1	−1
W Sicily versus E Sicily	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–
All Sicily versus All Italy	0.926	0.411	0.248	0.805	0.850	0.956	0.623	0.574
	−1	+1	+1	−1	−1	−1	−1	−1
S Spain versus Noncontact Iberia	0.337	0.841	–	0.872	0.679	1.000	0.519	0.981
	+1	−1	–	−1	−1	−1	+1	−1
Sardinia versus Italy	0.165	–	–	0.220	0.136	0.179	0.219	1.000
	+1	−1	–	+1	+1	+1	+1	−1
Sardinia versus Noncontact Iberia	0.053	1.000	–	0.069	0.045	0.037	0.104	1.000
	+1	−1	–	+1	+1	+1	+1	−1
Coastal Tunisia versus Inland Tunisia	–	0.177	0.006	0.021	–	–	0.001	0.304
	–	+1	+1	+1	+1	–	+1	+1
All Tunisia versus Morocco	1.000	0.059	0.000	0.868	1.000	1.000	0.245	0.039
	−1	+1	+1	−1	−1	−1	+1	+1
$\alpha = 0.05$	0.102	0.101	0.032	0.0022	0.012	0.012	0.0002	0.460
$\alpha = 0.30$	0.078	0.022	0.070	0.039	0.047	0.047	0.0017	0.986
Δf	0.033	0.193	0.063	0.387	0.274	0.377	0.019	0.997
Greek Test Sites								
Crete and Greece versus Cyprus	1.0000	0.9954	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000
	−1	−1	−1	−1	−1	−1	−1	+1
Crete and Greece versus Sicily	0.4261	0.0590	1.0000	0.3608	0.7137	0.9429	0.3181	0.0000
	+1	+1	−1	+1	−1	−1	+1	+1
Crete and Greece versus S Italy	0.734	0.046	–	0.490	0.731	0.869	0.188	0.000
	−1	+1	–	+1	−1	−1	+1	+1
S Italy versus N Italy	0.939	–	–	0.933	0.951	0.990	0.620	0.362
	−1	–	–	−1	−1	−1	−1	+1
Crete and Greece versus Malta	0.871	0.908	–	1.000	0.294	0.443	0.934	0.000
	−1	−1	–	−1	+1	+1	−1	+1
Crete and Greece versus Iberia	0.477	0.400	–	0.249	0.600	0.713	0.184	0.006
	+1	+1	–	+1	+1	−1	+1	+1
Turkey versus Phoenician Heartland and Periphery	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–
Turkey versus Syria	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–
Crete versus Phoenician Heartland and Periphery	0.815	0.390	1.0	0.862	0.939	0.949	0.759	0.000
	−1	+1	−1	−1	−1	−1	−1	+1
Crete versus Syria	0.272	0.618	–	0.367	0.754	0.349	0.408	0.000
	+1	−1	–	+1	−1	+1	+1	+1
Sicily versus Tunisia	0.014	0.605	1.000	0.749	0.053	0.004	0.906	0.426
	+1	−1	−1	−1	+1	+1	−1	+1
S Italy versus Coastal Tunisia	0.1255	1.0000	1.000	0.9171	0.4131	0.1802	0.9936	0.6772
	+1	−1	−1	−1	+1	+1	−1	−1
$\alpha = 0.05$	0.4013	0.3698	1.0000	1.0000	1.0000	0.4013	1.0000	0.0000
$\alpha = 0.30$	0.6172	0.8040	1.0000	0.9718	0.8507	0.9718	0.8507	0.0106
Δf	0.6230	0.7461	1.0000	0.8281	0.8281	0.8281	0.8281	0.0107

candidate but showed a strong score on the complementary Greek test matrix.

All of the PCS1+ through PCS3+ candidate central haplotypes are more than two steps away from each other, so the STR+s share no STR haplotypes. Therefore, their frequencies can be combined if sample counts are added together row by row to represent an aggregate PCS1+ through PCS3+ group. In general, across most geographical sites, the PCS1+, PCS2+, and PCS3+ groups combined to reinforce each other's Phoenician signals, boosting their aggregate scores but not their Greek scores (Table 3). The PCS1+, PCS2+, and PCS3+ frequencies in the Mediterranean region are represented in Figures 1D–1F.

PCS4+ through PCS6+ are all closely related to PCS1+ and PCS2+. Both PCS4+ and PCS5+ overlap PCS2+; PCS6+ does not, but shares a bridge PCS+ group (core 14,13,16,23,10,11,12) with both PCS1+ and PCS2+. Combining PCS4+ through PCS6+ with PCS1+ or PCS2+ would thus yield overcounting of some groups. Therefore, these are not included in the aggregate PCS1+ through PCS3+ group. It is notable that the range of STR+s in the cluster associated with PCS1+ and PCS2+ spans a range of five or six STR mutations, far in excess of that expected to emerge in the time since the Phoenician expansion. Although each STR+ covers geographically distinct colonies, each is rooted in the Phoenician heartland. This argues for a common source of related lineages rooted in Lebanon.

It can be deduced from the structure of the tests that admixture from other occupation of both Phoenician *non*-contact sites and contact sites would tend to systematically wash out the significance of a Phoenician signal throughout the range of the Phoenician Colony Specific test sites. For example, one of the five samples from Sardinia was PCS1+. Compared to Italy, at five out of 187, the probability of drawing this fraction by chance was 0.258, as reported in Table 3. If only 20% of the samples found in Italy were added to Sardinia's signal, this would have yielded two out of 47 for Sardinia, yielding a probability of 0.378, outside the $\alpha = 0.30$ threshold. Likewise, 30% of the Greek contribution of Crete in PCS1+ would raise the Fisher's exact probability from 0.173 to 0.328. The fact that this dilution did *not* systematically destroy a preponderant Phoenician signal argues that such admixture has been low enough to allow the isolation of components that were systematically Phoenician. The results presented here suggest that any additional expansions, such as the Jewish Diaspora, and subsequent population effects showed sufficiently low admixture or drift into both colonization sites and surrounding populations for a Phoenician signal to remain significant.

Haplogroup J2, in general, and haplotypes PCS1+ through PCS6+ therefore represent lineages that might have been spread by the Phoenicians; but could the patterns that we observe be accounted for by other events, particularly the Jewish Diaspora, for which we could not develop a formal test? Note that this is a separate question from that of whether they could *also* have been spread by other expansions: indeed, we expect that Jews of the Diaspora carried some of the same STR+ and SNP lineages with them as did Phoenicians of Phoenician expansion. Two lines of reasoning suggest to us that we must be detecting a distinct signal. First, the frequency of Jews in the Mediterranean region over almost all of our sample sites is currently less than 0.1%, and our own collection of samples contained no individuals who identified themselves as bearing Jewish heritage in a number of sites, such as Tunisia and Morocco.[22,23] Although historical admixture is expected to have occurred to some extent, recent studies tend to show strikingly low admixture in modern Jewish populations.[15] Second, any such admixture is likely to have contributed to both Phoenician contact and noncontact populations and thus could not explain a systematically *differential* signal. The excess of J2 (Figure 1B), PCS1+ (Figures 1C and 1D), PCS2+ (Figure 1E), and PCS3+ (Figure 1F) in coastal Tunisia, the site of Carthage, compared with inland Tunisia is particularly salient, because these lineages are considerably more rare in North Africa than in Southern Europe. It also suggests that the Roman destruction of Carthage did not eliminate the Carthaginian gene pool. Further support for the PCS+ haplotypes' spread with the Phoenicians is illustrated by their generally high frequency among the Phoenician contact sites across the Mediterranean basin (Figures 1D–1F).

The Greek expansion was not the focus of this study, but it nevertheless revealed several signals. In this case, two expansions from Western Europe that probably spread R1b chromosomes could potentially yield a "Greek" profile. According to Strabo, Brennus "the second" of the Prausi was attracted to Greece by internecine conflicts in 281 BCE. Subsequently, some of these people moved to Thracia in the north, with 20,000 of those moving to Galatia in the north-central Anatolian peninsula in 279–277 BCE.[24] Subsequent European genetic transfer occurred with the Crusades[16] and with European trade, leaving a general north-to-south gradient of R1b chromosomes, with a substantial concentration in Greece and Turkey, yielding a pattern that could resemble Greek colonization.

This study presents a methodology for constructing systematic tests identifying local signatures of colonization and for constructing aggregate scores measuring a consensus across all of the colonization sites. We have shown that the methodology does not produce significant signal for arbitrary sampling in noncolonization regions, and multiple markers that do not show patterns consistent with Phoenician colonization have been presented. Tests constructed to isolate Greek-colonization events from the Phoenician samples can show positive and weak scores both for Phoenician candidates and for non-Phoenician candidates, indicating that information is presented in those tests distinct from the Phoenician-colonization tests.

Application of this methodology to STR samples was more problematic as a result of prohibitively small samples at some sites and of nonuniform sample collection throughout the Mediterranean at this level of resolution, even when STR-only data were included. Smaller

collections limit the statistical power to resolve signals at any of the particular sites. With the possibility of single-step STR mutations in the intervening time allowed for, identification of candidate groups (STR+s) was possible. Although true mutated descendants will systematically augment the strength of the statistical resolution, this expansion of samples will also allow inclusion of identical-by-state haplotypes with distinct histories that might even derive from other haplogroups. In conclusion, there are many ways in which a colonization signal could be diluted to undetectable levels, but statistically robust signals should represent biologically meaningful events.

We do not suggest that the Phoenicians spread only or predominantly J2 and PCS1+ through PCS6+ lineages. They are likely to have spread many lineages from multiple haplogroups, but the lineages we highlight are the most highly differentiated ones providing the most readily detectable signals. Signals can only be detected when the same or related haplotypes were transmitted to multiple locations. Because most haplotypes are rare, the use of STR+s rather than individual haplotypes, and perhaps the preferential spread of a subset of pioneering or influential Phoenician families, might have enhanced our signal. The magnitude of the Phoenician contribution to the populations investigated was estimated from the candidate STR+'s prevalence in colony versus noncolony sites. Although the total fraction of colony samples contained within the PCS1+ through PCS3+ group is around 10%, it is the fraction above background, or the difference in frequencies between contact and noncontact sites (Table S4), that actually represents Phoenician signal. The mean difference in frequency was ~6%, providing a minimum estimate of the Phoenician input.

Given that these same lineages, including the STR haplotypes, were clearly spread in other ways as well, their identification in additional subjects would not in itself provide evidence that such people were of Phoenician descent. This, however, is a reflection of the limited phylogenetic resolution used, and it is reasonable to expect that future thorough searches for SNPs or STR combinations could lead to the discovery of rare but reliable markers of such descent. The technology for resequencing individual genomes at ever-decreasing cost makes this a realistic prospect.[24]

Finally, our work underscores the effectiveness of Y-chromosomal variability when combined with appropriate computational analysis for studying complex patterns of human migration, as well as the utility of wide geographical sampling with the use of a uniform marker set. This method is applicable to any type of genetic information from which descent could be inferred, such as mtDNA or autosomal regions with limited recombination, and within which enough markers are available to establish phylogeny. The numbers of sites passing at $\alpha = 0.3$ and $\alpha = 0.05$ levels to produce aggregates significant at the 5% level for various numbers of sites tested are outlined in Table 4. Therefore, even rather small sets at relatively low levels of significance can yield useful signal. Further applications could include systematic investigations of military expansions, such as the Greek signal, from the time of Alexander the Great in central and south Asia;[25] or the Mongol signal, carried through the military and trade-regulation activities to regions from China to Moscow[26] and south through North India, Iran, and Iraq. Trade and colonization without substantial military intervention also drove wealth and technological and cultural development. Examples of ways that genetic migration was mediated might include the silk and spice roads, which connected China with the Middle East through to Europe, as well as to spice sources in India and Indonesia, and the Incense Road, which connected India through the southern Arabian Peninsula. The Viking expansion involved not only raids but also significant trade and colonization, in multiple waves.[27] Important African centers of trade include Timbuktu, with archaeological evidence showing that Great Zimbabwe enjoyed goods from as far away as China. To complement investigations of known migrations, our methodology could also be used to search systematically for signals of expansion within a data set, starting from each site in turn, and could thus potentially discover unrecorded migrations as well.

Table 4. Number of Sites, *k*, with p Value Stronger than Significance Level α out of a Total of *N* Sites Tested that Are Required for the Aggregate to Pass at the 5% Significance Level

N	k ($\alpha = 0.30$)	k ($\alpha = 0.05$)
1	-	1
2	-	2
3	3	2
4	4	2
5	4	2
6	5	2
7	5	2
8	6	3
9	6	3
10	6	3
11	7	3
12	7	3
13	8	3
14	8	3
15	9	3

Supplemental Data

Supplemental Data include four tables and can be found with this paper online at http://www.ajhg.org/.

Acknowledgments

We thank the sample donors for taking part in this study, R. John Mitchell, Tad Schurr and other Genographic Project members for comments. We also thank Janet Ziegle and Applied Biosystems for genotyping support and Rabih Hosri for help with Figure 1. Y.X. and C.T.S. were supported by The Wellcome Trust, and M.A.J. by a Wellcome Trust Senior Fellowship in Basic Biomedical Science (057559). The Genographic Project is supported by

funding from the National Geographic Society, IBM and the Waitt Family Foundation.

Genographic Consortium members: Theodore G. Schurr (University of Pennsylvania, Philadelphia, PA, USA), Fabrício R. Santos (Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil), Lluis Quintana-Murci (Institut Pasteur, Paris, France), Jaume Bertranpetit (Universitat Pompeu Fabra, Barcelona, Catalonia, Spain), David Comas (Universitat Pompeu Fabra, Barcelona, Catalonia, Spain), Chris Tyler-Smith (The Wellcome Trust Sanger Institute, Hinxton, UK), Pierre A. Zalloua (Lebanese American University, Chouran, Beirut, Lebanon), Elena Balanovska (Russian Academy of Medical Sciences, Moscow, Russia), Oleg Balanovsky (Russian Academy of Medical Sciences, Moscow, Russia), R. John Mitchell (La Trobe University, Melbourne, Victoria, Australia), Li Jin (Fudan University, Shanghai, China), Himla Soodyall (National Health Laboratory Service, Johannesburg, South Africa), Ramasamy Pitchappan (Madurai Kamaraj University, Madurai, Tamil Nadu, India), Alan Cooper (University of Adelaide, South Australia, Australia), Lisa Matisoo-Smith (University of Auckland, Auckland, New Zealand), Ajay K. Royyuru (IBM, Yorktown Heights, New York, USA), Daniel E. Platt (IBM, Yorktown Heights, New York, USA), Laxmi Parida (IBM, Yorktown Heights, New York, USA), Jason Blue-Smith (National Geographic Society, Washington, D.C., USA), David F. Soria Hernanz (National Geographic Society, Washington, D.C., USA), and R. Spencer Wells (National Geographic Society, Washington, D.C., USA).

Received: September 11, 2008
Revised: October 11, 2008
Accepted: October 14, 2008
Published online: October 30, 2008

References

1. Stieglitz, R.R. (1990). The geopolitics of the Phoenician Littoral in the Early Iron Age. Bull. Am. Schools Orient. Res. *279*, 9–12.
2. Moscati, S. (1973). The World of Phoenicians (London: Weidenfeld and Nicolson Ltd.).
3. Markoe, G. (2000). Phoenicians (London: British Museum Press).
4. Aubet, M.E. (2001). The Phoenicians and the West: Politics, Colonies and Trade (Cambridge: Cambridge University Press).
5. Marston, E. (2002). The Phoenicians (New York: Benchmark Books).
6. Harden, D. (1971). The Phoenicians (London: Penguin Books).
7. Jobling, M.A., and Tyler-Smith, C. (2003). The human Y chromosome: an evolutionary marker comes of age. Nat. Rev. Genet. *4*, 598–612.
8. Rosser, Z.H., Zerjal, T., Hurles, M.E., Adojaan, M., Alavantic, D., Amorim, A., Amos, W., Armenteros, M., Arroyo, E., Barbujani, G., et al. (2000). Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am. J. Hum. Genet. *67*, 1526–1543.
9. Semino, O., Magri, C., Benuzzi, G., Lin, A.A., Al-Zahery, N., Battaglia, V., Maccioni, L., Triantaphyllidis, C., Shen, P., Oefner, P.J., et al. (2004). Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the Neolithization of Europe and later migratory events in the Mediterranean area. Am. J. Hum. Genet. *74*, 1023–1034.
10. Semino, O., Passarino, G., Brega, A., Fellous, M., and Santachiara-Benerecetti, A.S. (1996). A view of the Neolithic demic diffusion in Europe through two Y chromosome-specific markers. Am. J. Hum. Genet. *59*, 964–968.
11. Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., et al. (2000). The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. Science *290*, 1155–1159.
12. Mitchell, R.J., Earl, L., and Fricke, B. (1997). Y-chromosome specific alleles and haplotypes in European and Asian populations: linkage disequilibrium and geographic diversity. Am. J. Phys. Anthropol. *104*, 167–176.
13. Tripolitis, A. (2001). Religions of the Hellenistic-Roman age (Michigan: Wm. B. Eerdmans Publishing Company).
14. Barnavi E., ed. (1994). A Historical Atlas of the Jewish People: From the Time of the Patriarchs to the Present (New York: Schocken).
15. Behar, D.M., Garrigan, D., Kaplan, M.E., Mobasher, Z., Rosengarten, D., Karafet, T.M., Quintana-Murci, L., Ostrer, H., Skorecki, K., and Hammer, M.F. (2004). Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations. Hum. Genet. *114*, 354–365.
16. Zalloua, P.A., Xue, Y., Khalife, J., Makhoul, N., Debiane, L., Platt, D.E., Royyuru, A.K., Herrera, R.J., Hernanz, D.F., Blue-Smith, J., et al. (2008). Y-chromosomal diversity in Lebanon is structured by recent historical events. Am. J. Hum. Genet. *82*, 873–882.
17. Schlecht, J., Kaplan, M.E., Barnard, K., Karafet, T., Hammer, M.F., and Merchant, N.C. (2008). Machine-learning approaches for classifying haplogroup from Y chromosome STR data. PLoS Comput Biol *4*, e1000093.
18. Agrawal, R., and Srikant, R. (1994). Fast Algorithms for Mining Association Rules. Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94), pp. 487–499.
19. Arredi, B., Poloni, E.S., Paracchini, S., Zerjal, T., Fathallah, D.M., Makrelouf, M., Pascali, V.L., Novelletto, A., and Tyler-Smith, C. (2004). A predominantly Neolithic origin for Y-chromosomal DNA variation in North Africa. Am. J. Hum. Genet. *75*, 338–345.
20. Sokal, R.R., and Rohlf, F.J. (1995). Biometry: The Principles and Practice of Statistics in Biological Research (New York: W. H. Freeman).
21. Zhivotovsky, L.A., Underhill, P.A., Cinnioglu, C., Kayser, M., Morar, B., Kivisild, T., Scozzari, R., Cruciani, F., Destro-Bisol, G., Spedini, G., et al. (2004). The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. Am. J. Hum. Genet. *74*, 50–61.
22. Hanford J.V., ed. (2004). 2004 International Religious Freedom Report (Washington: U.S. Department of State).
23. DelaPergola, S. (2002). World Jewish Population 2002. In American Jewish Yearbook 2002, D. Singer and L. Grossman, eds. (NY: American Jewish Committee), pp. 247–274.
24. Mardis, E.R. (2008). The impact of next-generation sequencing technology on genetics. Trends Genet. *24*, 133–141.
25. Firasat, S., Khaliq, S., Mohyuddin, A., Papaioannou, M., Tyler-Smith, C., Underhill, P.A., and Ayub, Q. (2007). Y-chromosomal evidence for a limited Greek contribution to the Pathan population of Pakistan. Eur. J. Hum. Genet. *15*, 121–126.
26. Zerjal, T., Xue, Y., Bertorelle, G., Wells, R.S., Bao, W., Zhu, S., Qamar, R., Ayub, Q., Mohyuddin, A., Fu, S., et al. (2003). The genetic legacy of the Mongols. Am. J. Hum. Genet. *72*, 717–721.
27. Bowden, G.R., Balaresque, P., King, T.E., Hansen, Z., Lee, A.C., Pergl-Wilson, G., Hurley, E., Roberts, S.J., Waite, P., Jesch, J., et al. (2008). Excavating past population structures by surname-based sampling: the genetic legacy of the Vikings in northwest England. Mol. Biol. Evol. *25*, 301–309.

LI

APPENDIX I: FUNERALE DI VENEZIA

Venessia.com

Funerale di Venezia

Com.: 003/2009

Venezia, 7 novembre 2009

RICERCA SULLE ORIGINI DEI VENEZIANI ATTRAVERSO LO STUDIO DEL DNA

COLLABORAZIONE TRA WORCESTER POLYTECHNIC INSTITUTE E NATIONAL GEOGRAPHIC SOCIETY

In occasione del Funerale di Venezia verranno effettuati, da studenti e ricercatori dello Worcester Polytechnic Institute del Massachusetts (USA), prelievi non invasivi di campioni di saliva, su chi vorrà sottoporsi all'esperimento, che sono parte di una ricerca che si prefigge di raccogliere dati genetici per chiarire le origini delle popolazioni dell'Europa centro-occidentale.
Lo studio è infatti inserito nel progetto internazionale "The Genographic Project: Molecular Genetic Analyses of Western/Central European Populations" sponsorizzato dall'autorevole National Geographic Society che si prefigge di raccogliere campioni genetici su oltre 100mila individui in tutto il mondo.
È condotto dal Venice Project Center del Worcester Polytechnic Institute e, scrivono i promotori, «si concentra sul patrimonio genetico dei Veneti in generale e dei Veneziani in particolare, con l'intento di contribuire ad una migliore comprensione delle origine di Venezia. I dati raccolti serviranno a confermare o smentire diverse teorie sulle origini geografiche dei Veneti, sulla loro diffusione nel continente europeo e soprattutto sul loro insediamento nelle lagune dove oggi sorge Venezia».

Dal punto di vista tecnico è piuttosto semplice sottoporsi all'esperimento che consiste nel prelievo di un campione di saliva con un tampone a forma di spazzolino da denti con un'estremità molto simile a un cotton-fiocc. Sarà lo stesso volontario a mettere in bocca il tampone, poggiandolo sulla parte interna della guancia, dove dovrà essere tenuto per circa due minuti. Al volontario saranno inoltre richieste semplici informazioni sull'origine della propria famiglia e verrà richiesto di firmare un consenso alla raccolta del campione.

I campioni, raccolti nell'assoluto rispetto della privacy del volontario e dei suoi antenati, non verranno utilizzati per ottenere informazioni medico-sanitarie sui soggetti che si sono sottoposti all'esperimento. Verranno custoditi presso il Centro di Evoluzionistica dell'Università "Pompeu Fabra" di Barcellona. A garanzia della privacy ogni campione verrà associato a un codice segreto conosciuto solo dal volontario e dai responsabili centrali del progetto, e non dagli operatori che raccolgono i campioni. In ogni caso in qualunque momento il volontario potrà chiedere la distruzione del campione a lui associato.

Pierluigi TAMBURRINI – Addetto Stampa Funerale di Venezia
Tel. 349.6474663 – E-Mail: p.tamburrini@gmail.com - stampa@venessia.com

APPENDIX J: PHOTOS FROM TESTING

FIGURE 44: PRESS CONFERENCE TESTING

FIGURE 45: FUNERAL FOR VENICE TESTING